Search for the SM Higgs Boson decaying to $b\bar{b}$ and $\tau\tau$ at CMS

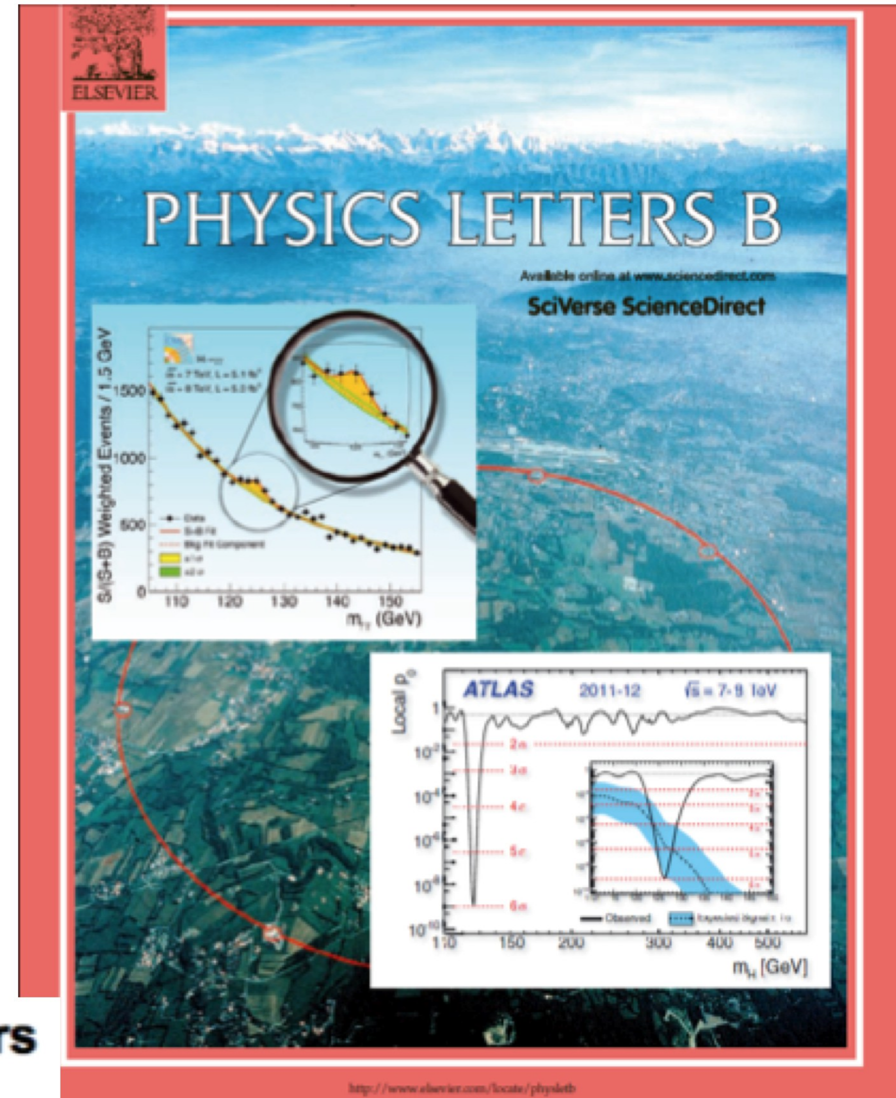
David Lopes Pegna (Princeton University-LPC FNAL)
On behalf of the CMS Collaboration

Higgs Workshop, Brookhaven
1 October 2012

July 4th: an historical event

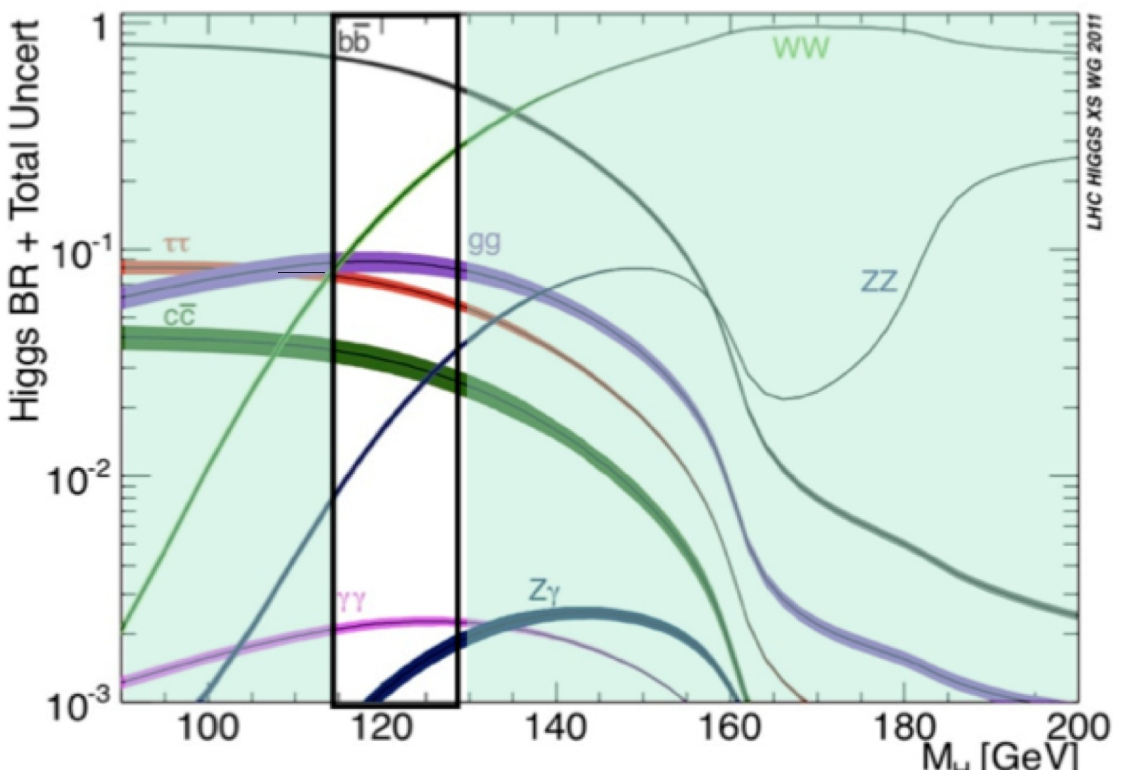
@ CERN

@ Melbourne



'God Particle' 'Discovered': European Researchers Claim Discovery of Higgs Boson-Like Particle

If it is the Higgs....



@125 GeV:

BR(H → bb) ~ 58%

BR(H → WW) ~ 22%

BR(H → $\tau\tau$) ~ 6%

BR(H → ZZ*) ~ 3%

BR(H-> γγ) ~0.22%

Our goal now is to confirm or exclude it's the Standard Model Higgs

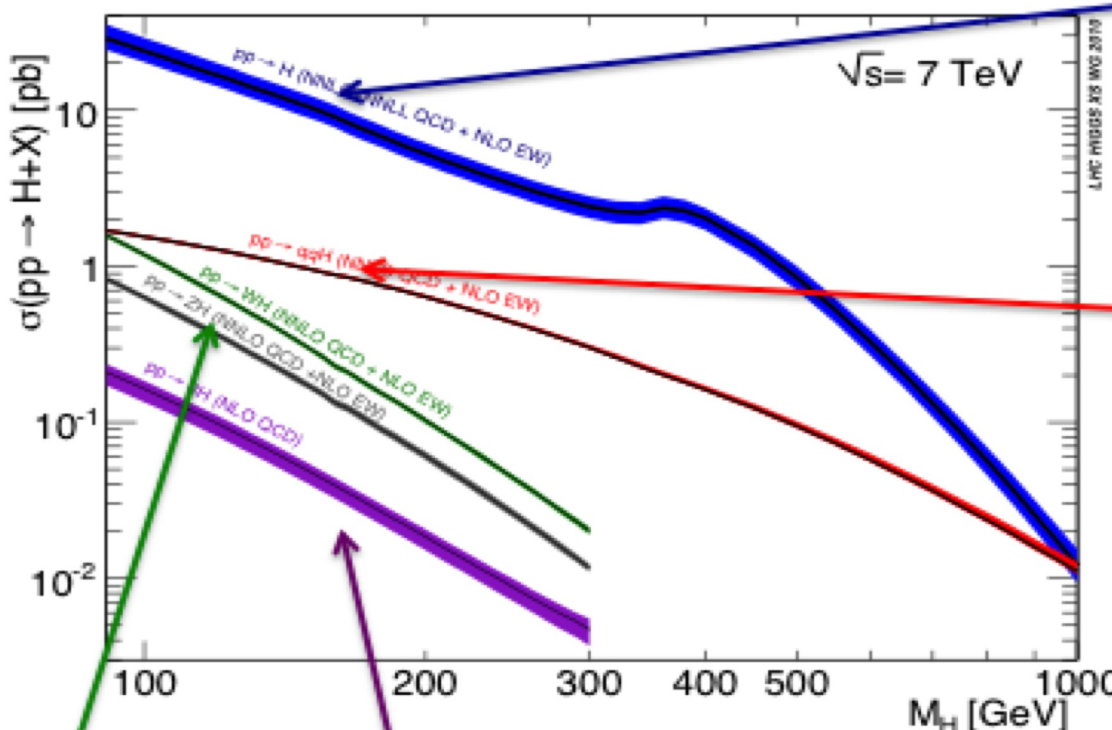
→ need complementary information from as many channels as possible

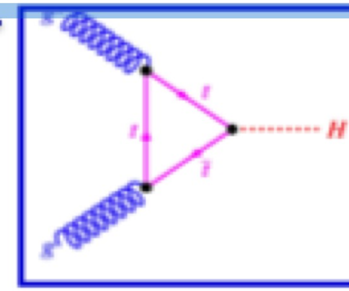
→ H → bb largest Branching Ratio by far below 130 GeV

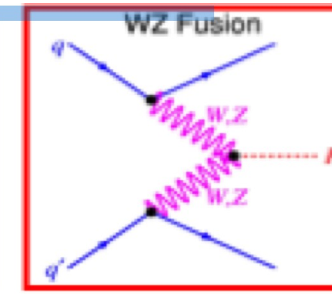
→ BR(H → gg) + BR(H → cc) ~13%, w/o H → bb, ¾ of the width would be invisible!

→ H → $\tau\tau$: crucial information on lepton coupling (could it be leptophobic?)

Production Modes







gg fusion

Vector Boson Fusion

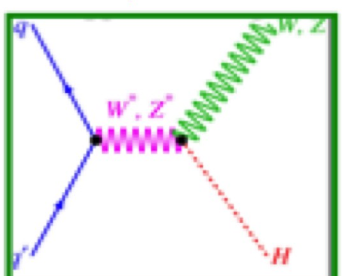
H → ττ

Can be studies in all production modes

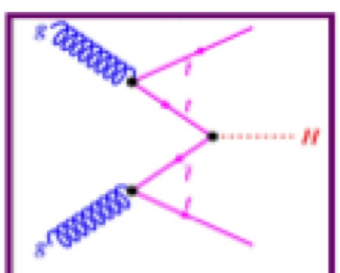
gg fusion challenging

VBF cleaner, VH recently added

Higgs-strahlung: ZH,WH



Associated production With ttbar pair

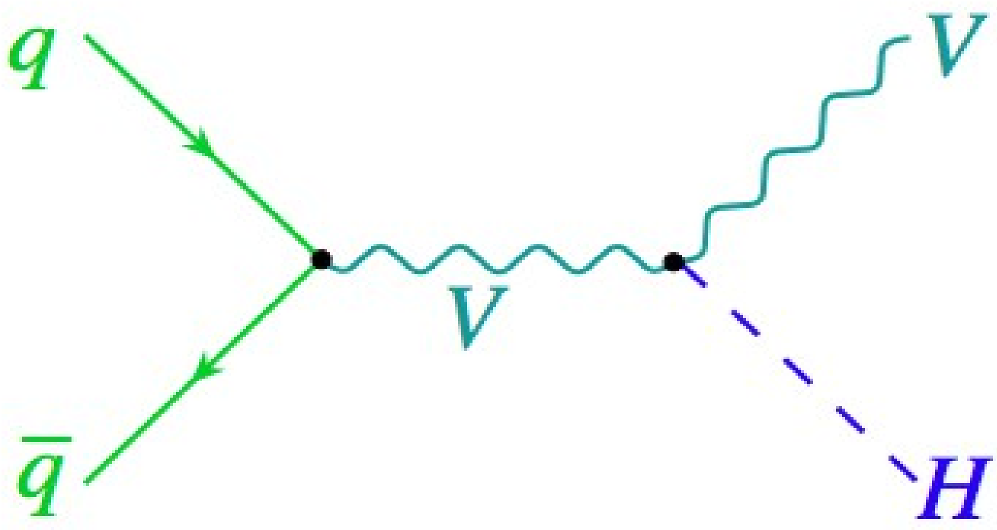


H → bb

Overwhelming QCD background makes gg → H → bb impossible

Look at associated production

ttH: even less rate, but accessible with large ∫Ldt

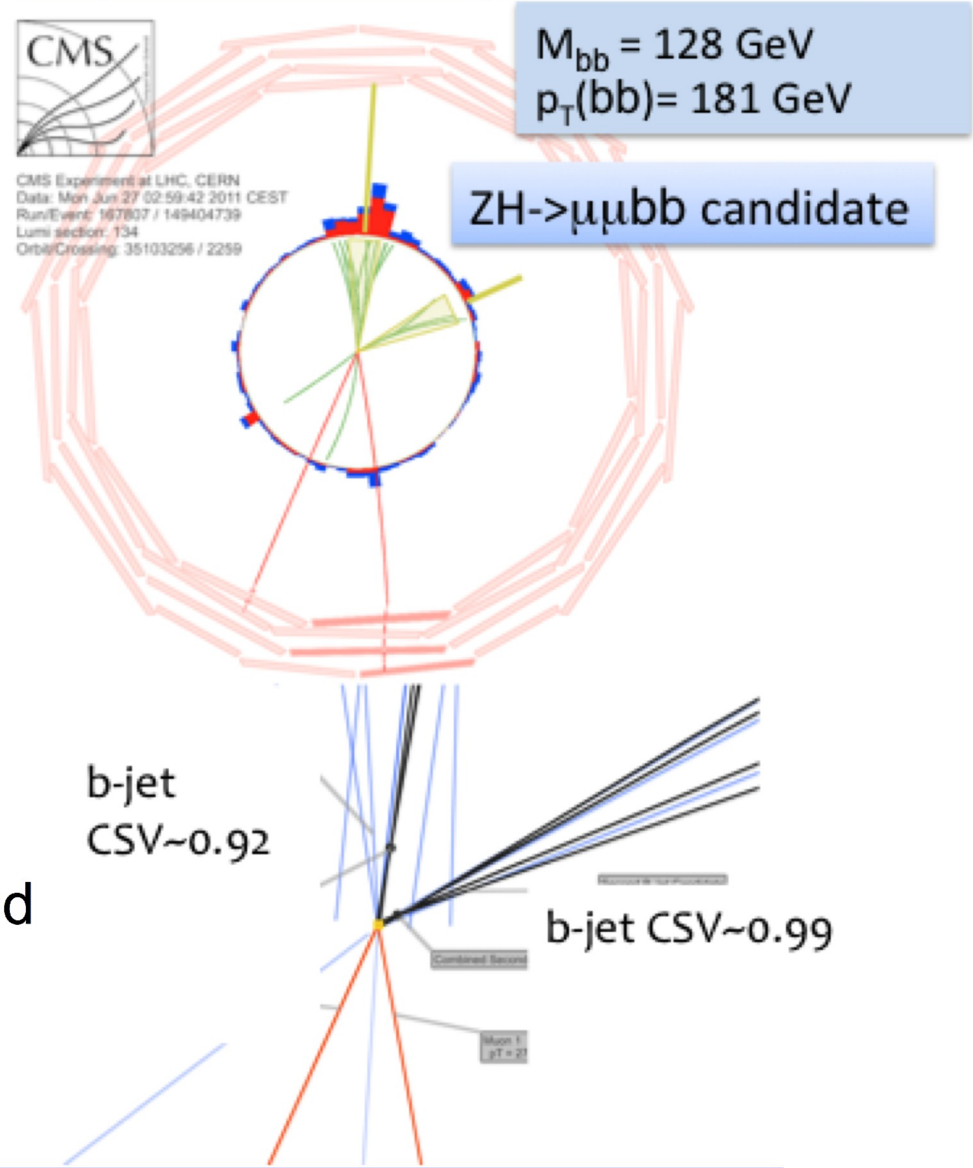


$VH \rightarrow Vb\bar{b}$
$V \rightarrow ll,\ l\nu,\ \nu\nu$

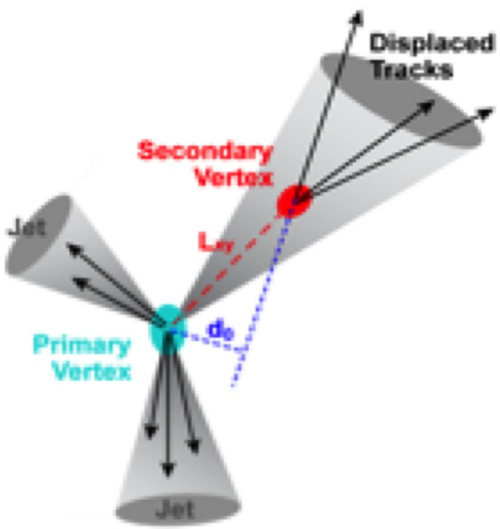
Most sensitive channel with b in final states
Intriguing excess in the Tevatron VH → bb analysis

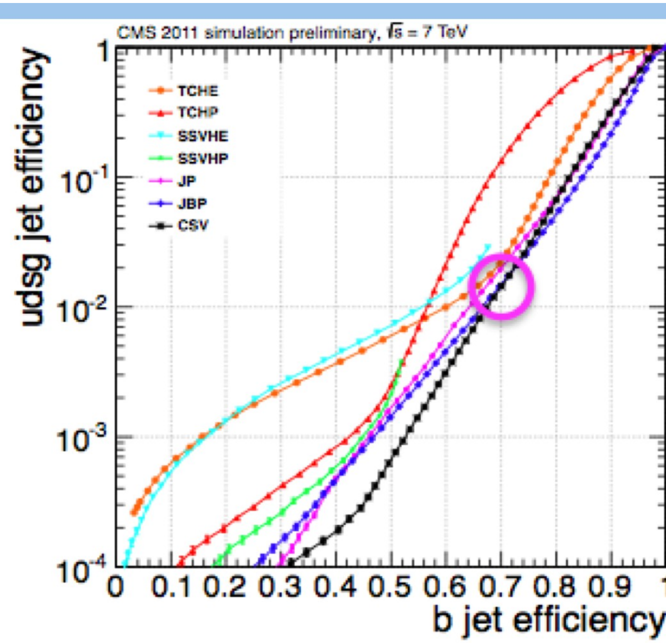
VH Analysis in a nutshell

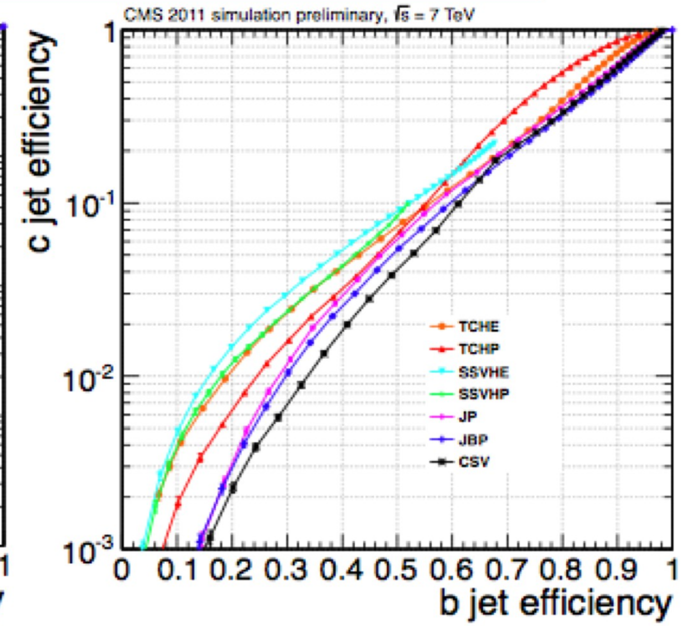
- First CMS Vhbb analysis on 7 TeV data: Phys. Lett. B 710(2012) 284-306
- 5 modes under study:
  Z(ll)H, W(lν)H, Z(νν)H, l= e,μ
- Boosted analysis (better S/B):
  → Require high momentum vector
  boson and 2-b tagged jet,
  back-to -back
- Use Data control regions to constrain
  most important backgrounds
  (V+jet, Light or Heavy, ttbar)
- Boosted Decision Tree algorithm (BDT)
  to discriminate signal versus background
- Improvements since 2011:
  → b-jet energy regression
  → Two p_T(V) bins



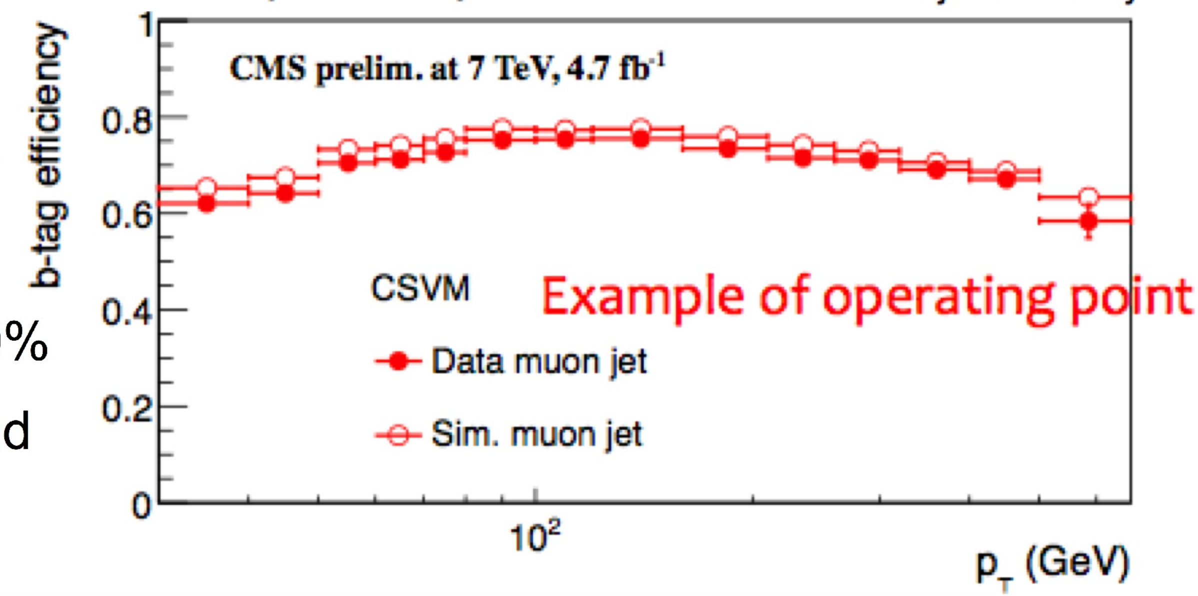
b-tagging at CMS







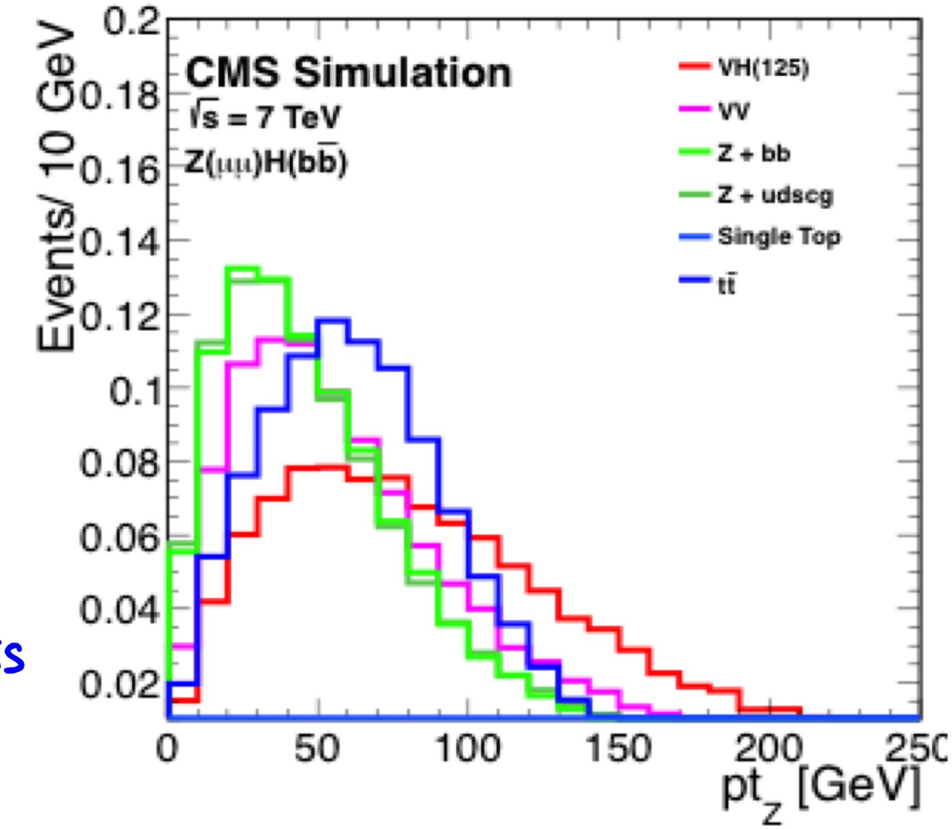
- CSV: Likelihood tagger using SV (if any), track IP etc.
- Eff ~70% for udsg ~2%, c-jet eff~20%
- Eff and fake rate from data: ttbar and μ+jet events



Event Categories

- Boost topology requirement is the name of the game
  → original proposal by Butterworth et al. in 2008 in the context of substructure analysis

- Split events in two categories based on $p_T(V)$
  New since 2011 Analysis
  → increase acceptance in lower boost region, backgrounds still manageable
  → Lower threshold possible in Z(ll)H due to additional ttbar suppression

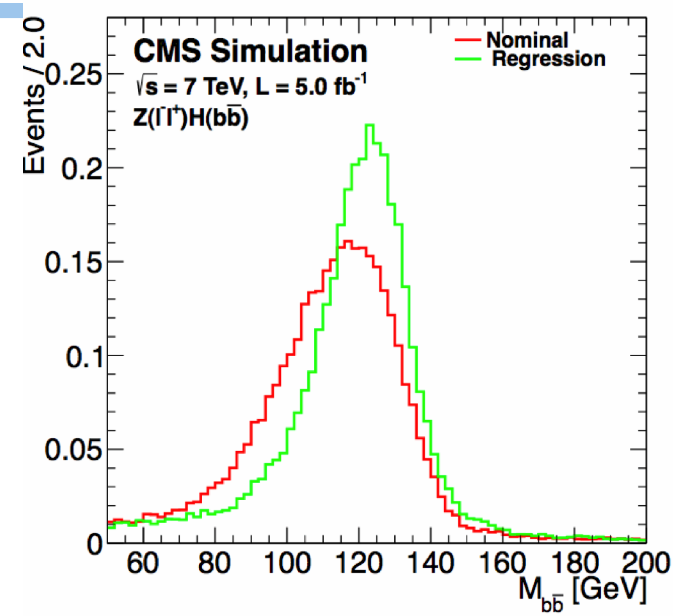


Channel	Medium boost	High boost
ZllH	50<Zpt<100	Zpt>100
WlnH	120< Wpt<170	Wpt>170
ZnnH	120<Zpt<160	Zpt>160

B-jet energy Regression

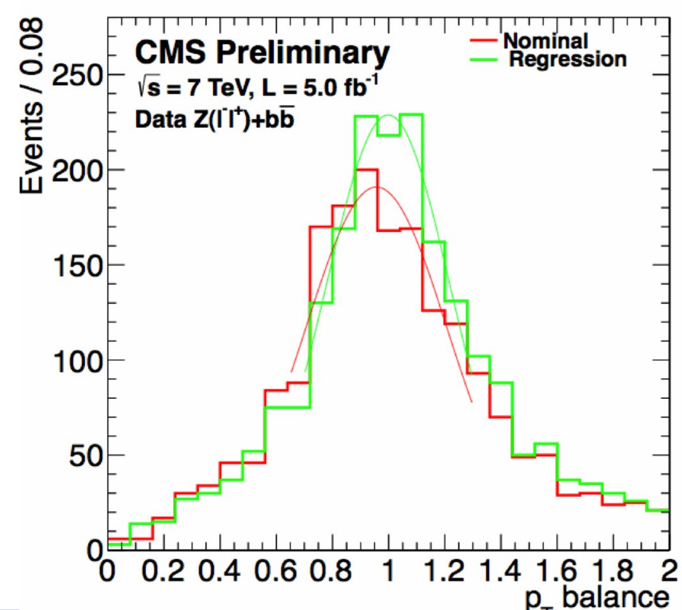
- Implementation based on NN method developed at CDF for b-jet energy corrections: http://arxiv.org/pdf/1107.3026.pdf

New since 2011 Analysis



- Multivariate Regression (BDT) trained on VH signal events using several (b)-jet variables

- Improvements in resolution of the order of 20% for Z(ll)H, 15% for W($l\nu$)H and Z($\nu\nu$)

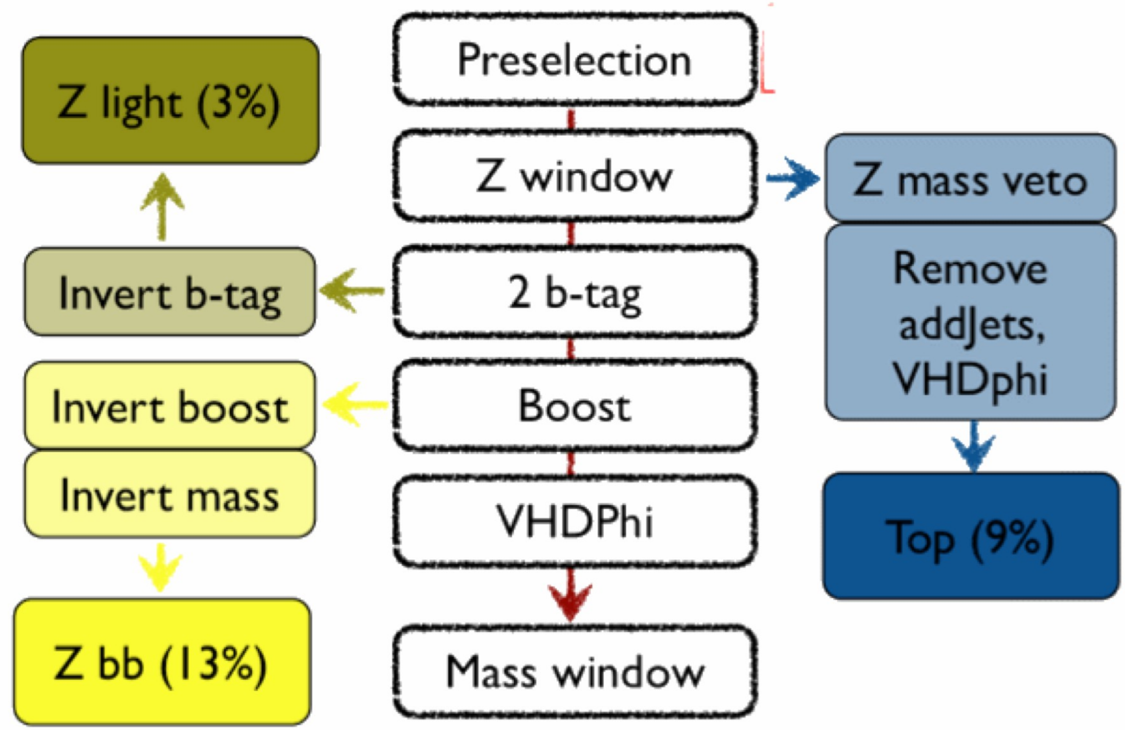


- Extensively validated on simulation and Data Control Regions (Z(ll)+bb, ttbar, Single Top)

Background Control Regions

- Define several CRs enriched in different background components
- Kinematic selection as close as possible to the one for the Signal Region (SR)
- Scale Factors (SF) for V+light jets, ttbar and V+heavy jets determined simultaneously in each mode from simultaneous binned Maximum Likelihood fit

Example: Zee control region definition



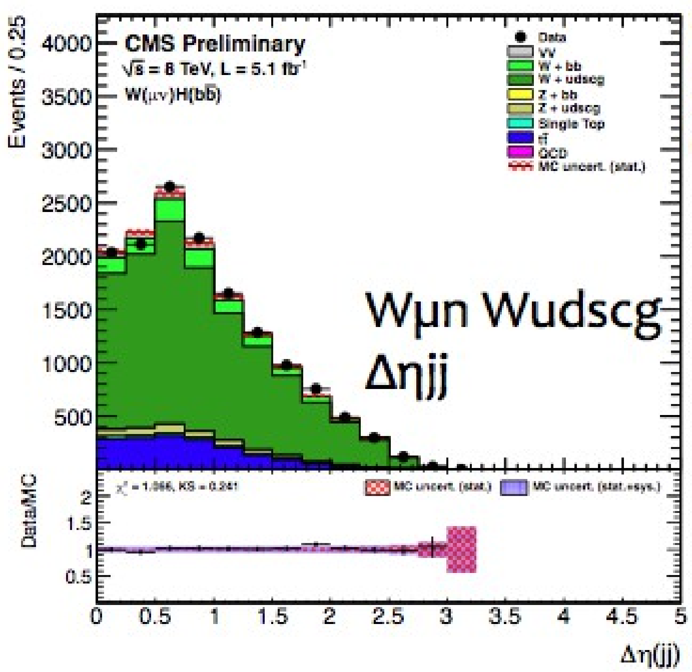
New since 2011 Analysis

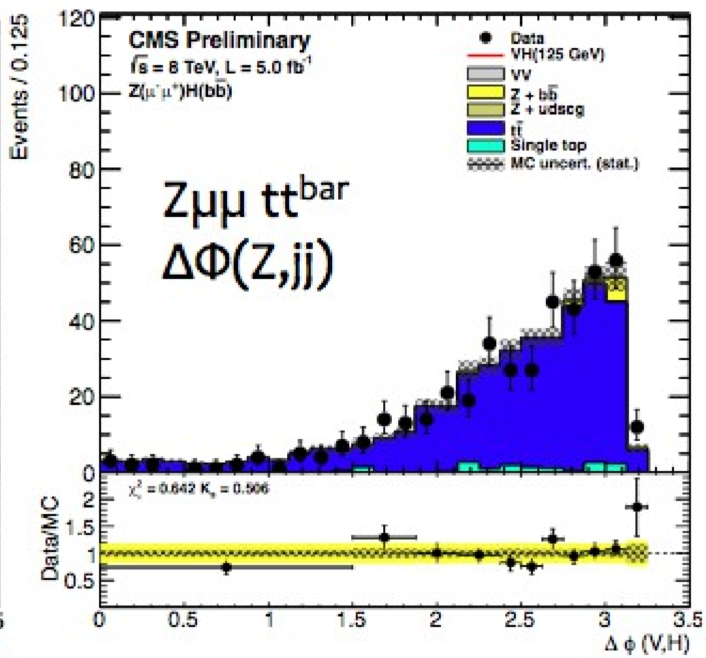
- Renormalize background estimates in Signal region based on Scale Factors: **$B(SR) = SF(CR) * B_{MC}(SR)$**

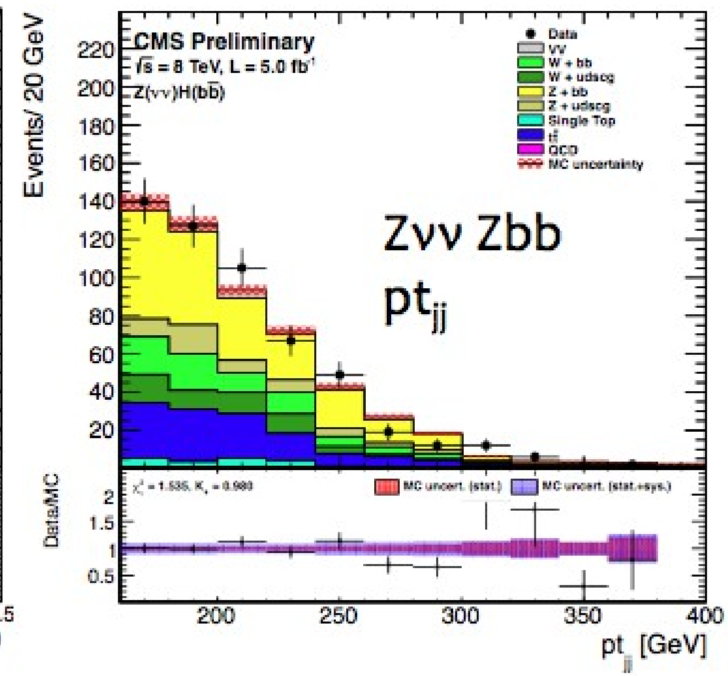
Background Control Regions

- Example of data/MC agreement in the Control Regions for variables used in the analysis
  → Many more in backup

- Calibrate most important backgrounds, test analysis robustness







BDT: Event Selection

- Preselection cuts on:
  → boost topology
  → b-tag enriched

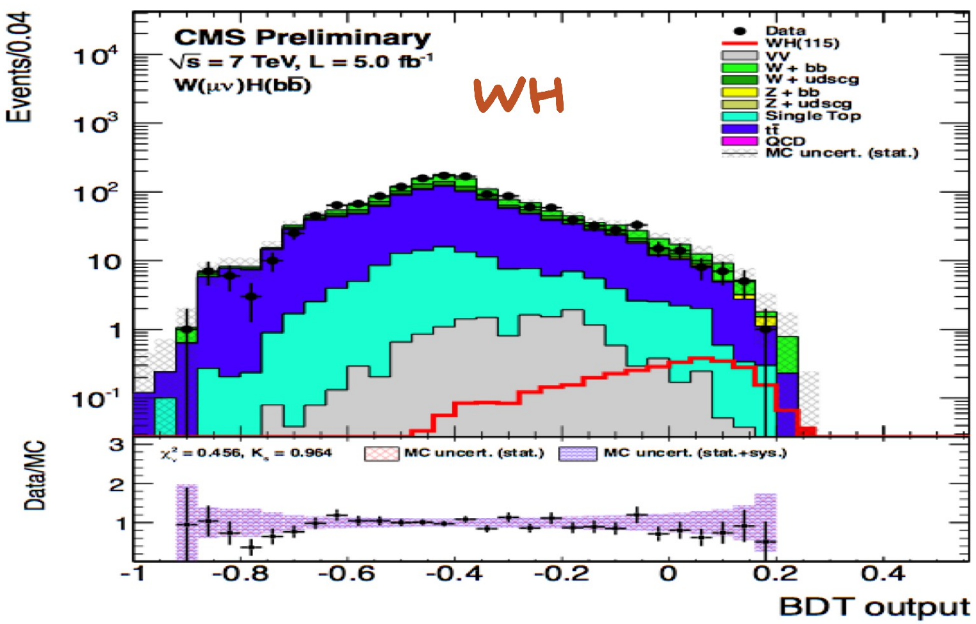
- Set of variables in the BDT largely overlapping with 2011 analysis

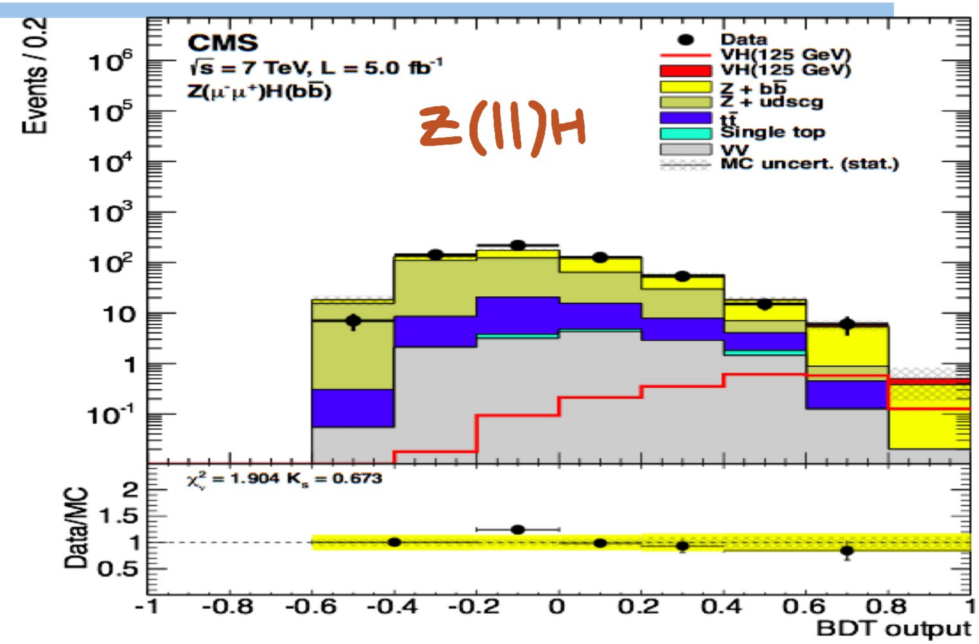
Variable	W($\ell\nu$)H	Z($\ell\ell$)H	Z($\nu\nu$)H
$m_{\ell\ell}$	–	$75 < m_{\ell\ell} < 105$	–
$p_T(j_1)$	> 30	> 20	> 80
$p_T(j_2)$	> 30	> 20	> 20
$p_T(jj)$	> 120	–	$120 - 160$ (> 160)
$m(jj)$	< 250	$80 < m(jj) < 150$ (–)	< 250
$p_T(V)$	$120 - 170$ (> 170)	$50 - 100$ (> 100)	–
CSV_{max}	> 0.40	0.50 (0.244)	> 0.50
CSV_{min}	> 0.40	0.244	> 0.50
N_{al}	$= 0$	–	$= 0$
$\Delta\phi(E_T^{miss}, jet)$	–	–	> 0.5
E_T^{miss}	> 35 (elec)	–	$120 - 160$ (> 160)
BDT	full distribution	full distribution	full distribution

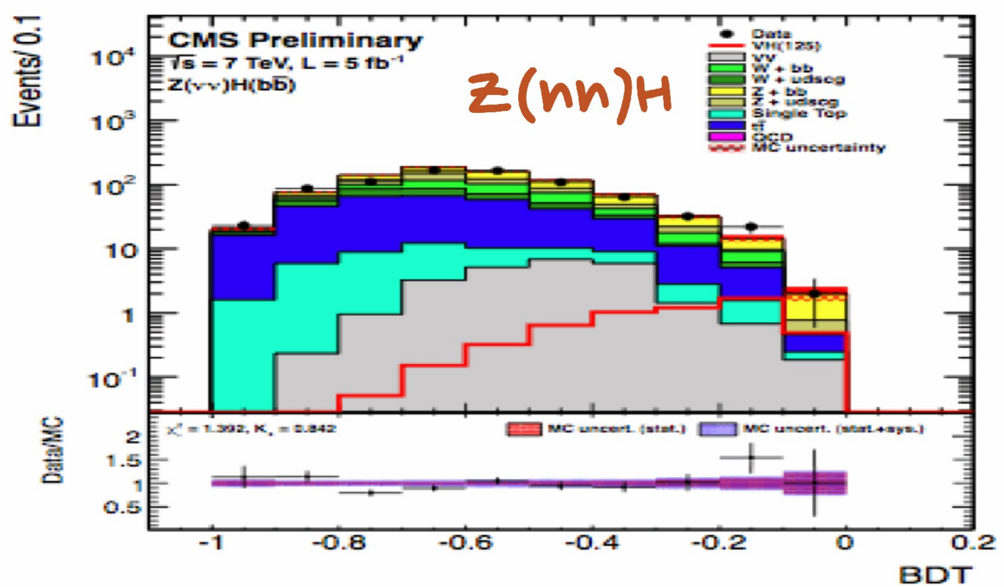
Variable
p_{Tj}: transverse momentum of each Higgs daughter
$M(jj)$: dijet invariant mass
$p_T(jj)$: dijet transverse momentum
$p_T(V)$: vector boson transverse momentum (or pfMET)
CSV1: value of CSV for best b-tagged jet
CSV2: value of CSV for second-best b-tagged jet
$\Delta\phi(V,H)$: azimuthal angle between V (or pfMET) and dijet
$\Delta\eta(J1,J2)$; difference in η between Higgs daughters
$\Delta R(J1,J2)$; distance in η–ϕ between Higgs daughters
N_{aj}: number of additional jets ($p_T > 30$ GeV, $\lvert\eta\rvert < 4.5$)
$\Delta\phi(pfMET, J)$(only for $Z(\nu\nu)H$)
$\Delta\theta_{pull}$: color pull angle

- Limit extraction based on shape analysis on BDT output: About 20% improvement in expected limit w.r.t. 2011 Cut and count in Signal enriched region

BDT Analysis (7 TeV)

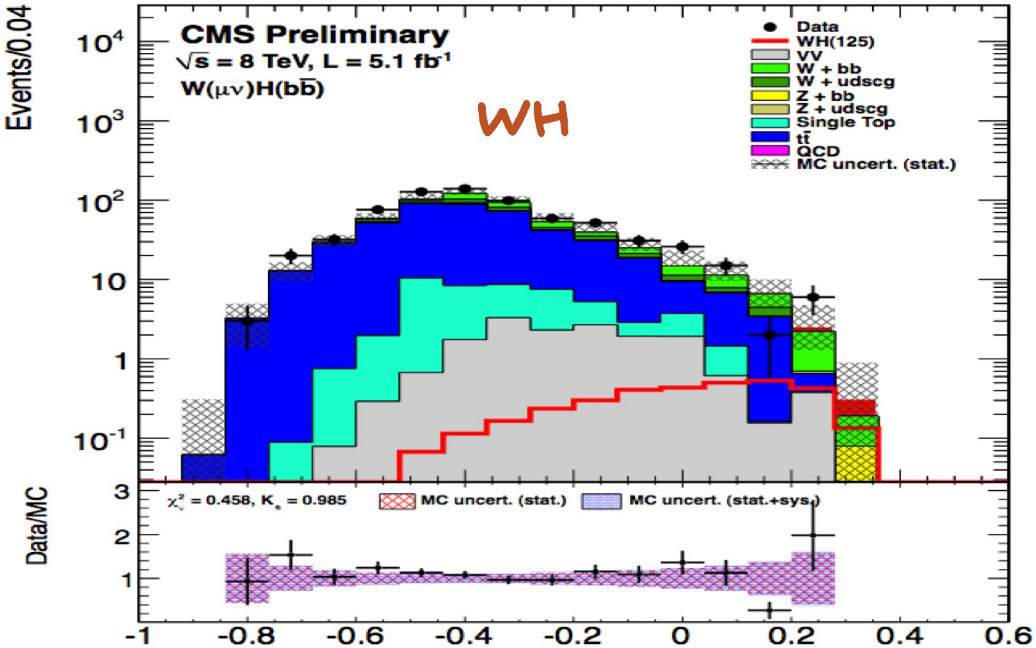


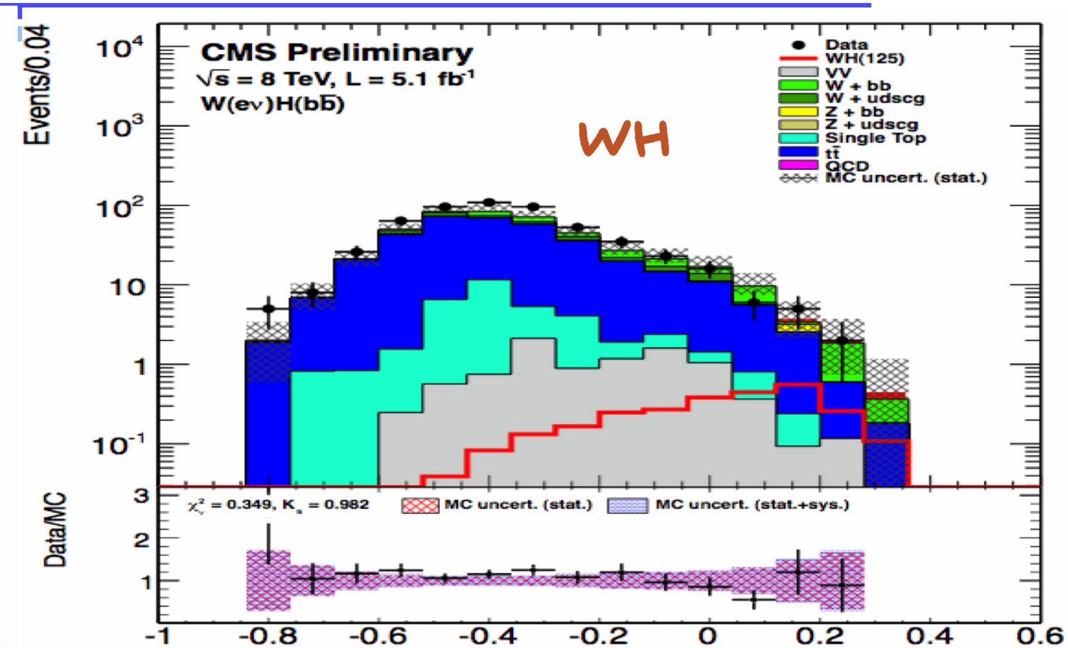


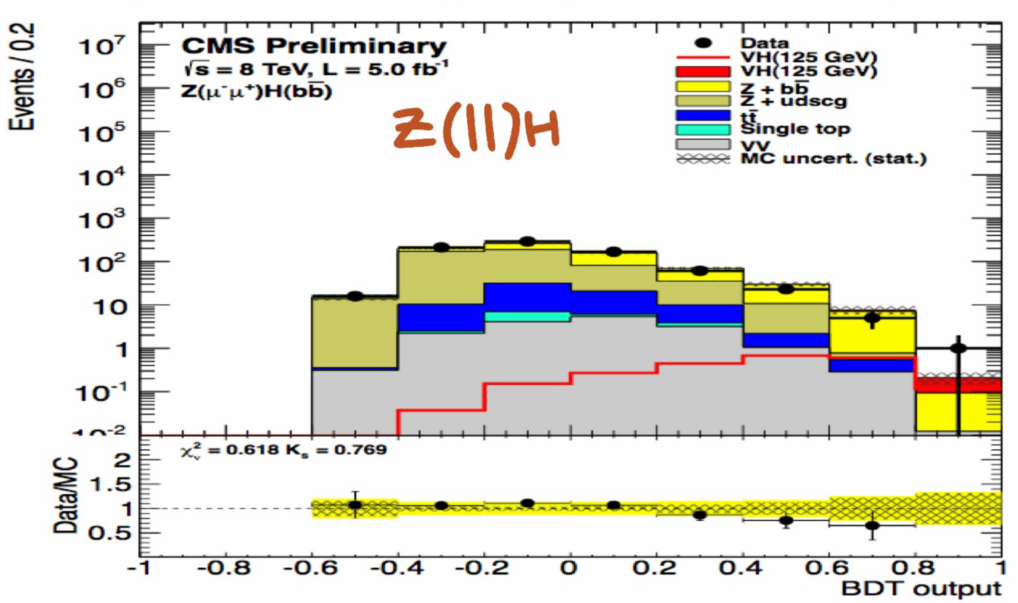


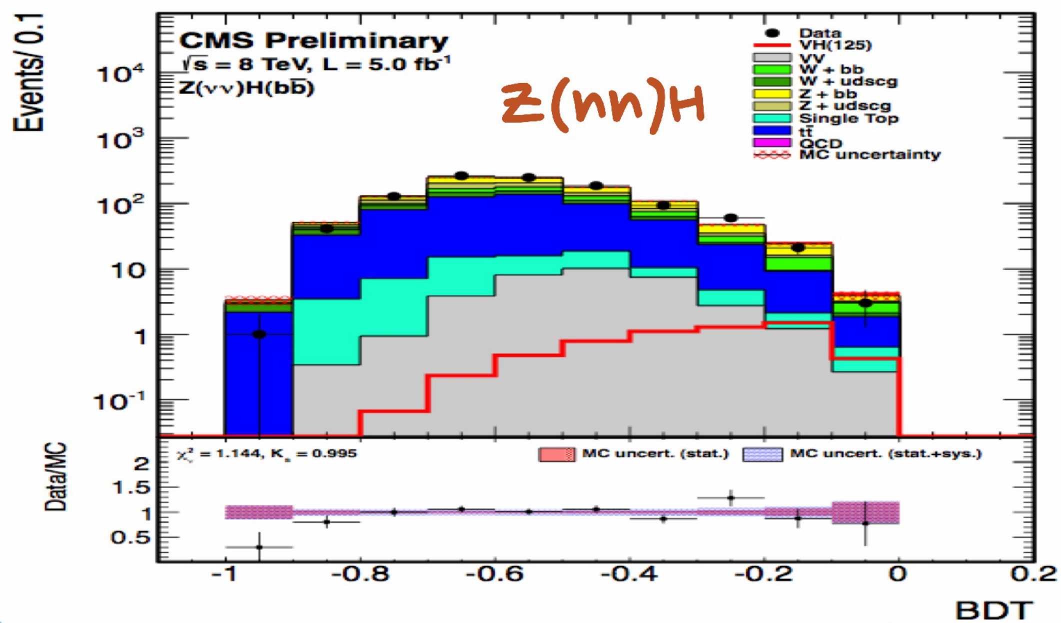
Analyses at 7 and 8 TeV carried on separately, final results from combination of:
5 (channels) x 2 (p_T bins) x 2 (7+8 TeV) =20 BDT discriminant fits at each m_H (110-135 GeV)

BDT Analysis (8 TeV)







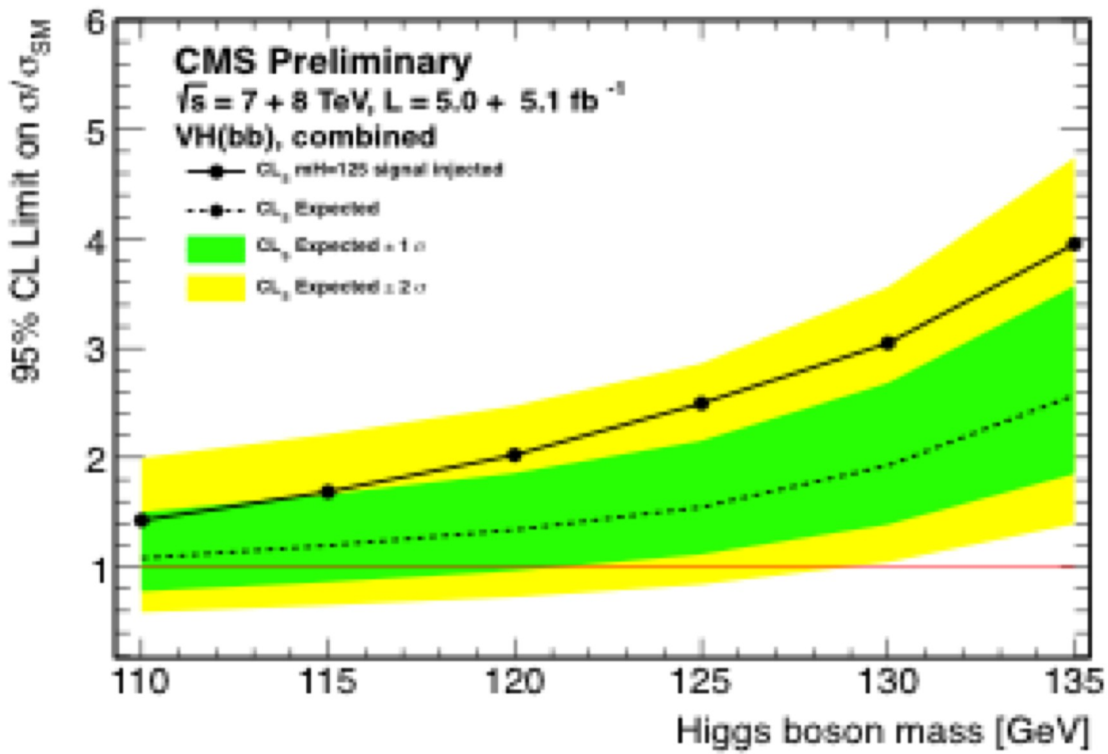


Systematic Uncertainties

Source	Range
Luminosity	2.2-4.4%
Lepton efficiency and trigger (per lepton)	3%
$Z(\nu\nu)H$ triggers	2%
Jet energy scale	2–3%
Jet energy resolution	3–6%
Missing transverse energy	3%
b-tagging	3–15%
Signal cross section (scale and PDF)	4%
Signal cross section (p_T boost, EWK/QCD)	5–10% / 10%
Signal Monte Carlo statistics	1-5%
Backgrounds (data estimate)	$\approx$ 10%
Diboson and single-top (simulation estimate)	30%

Dominant uncertainties: b-tagging, background modeling, signal cross-section

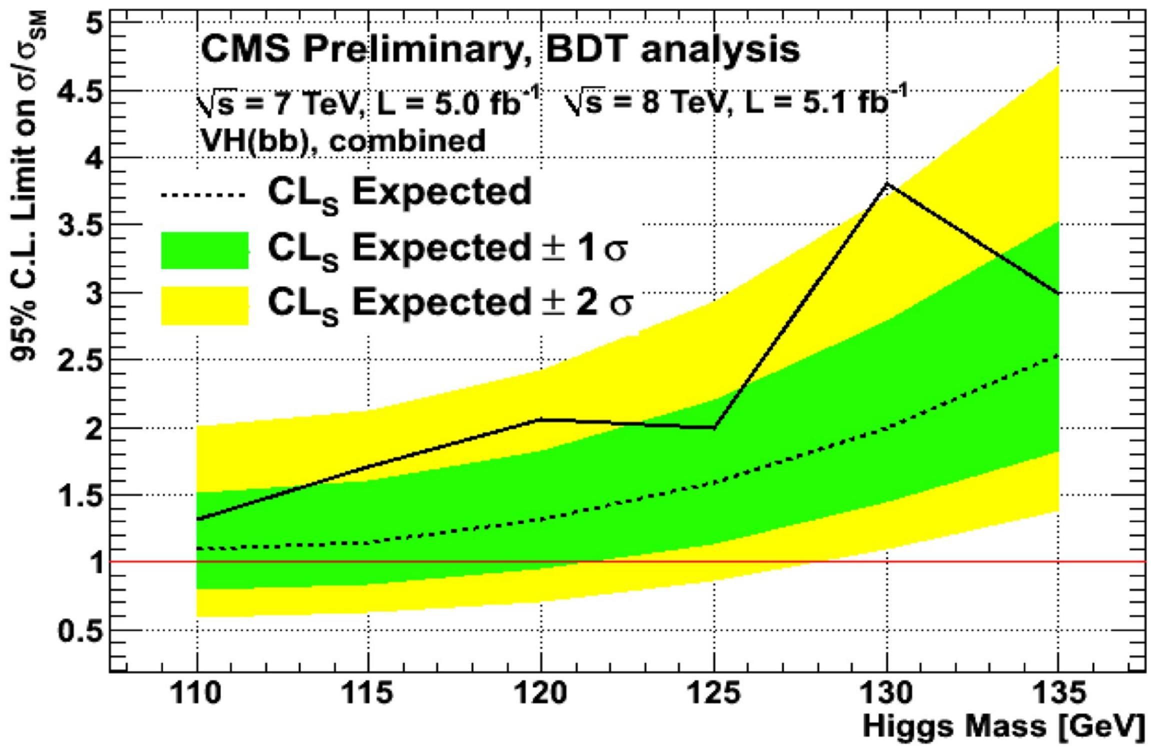
Results: SM Exclusion Limits

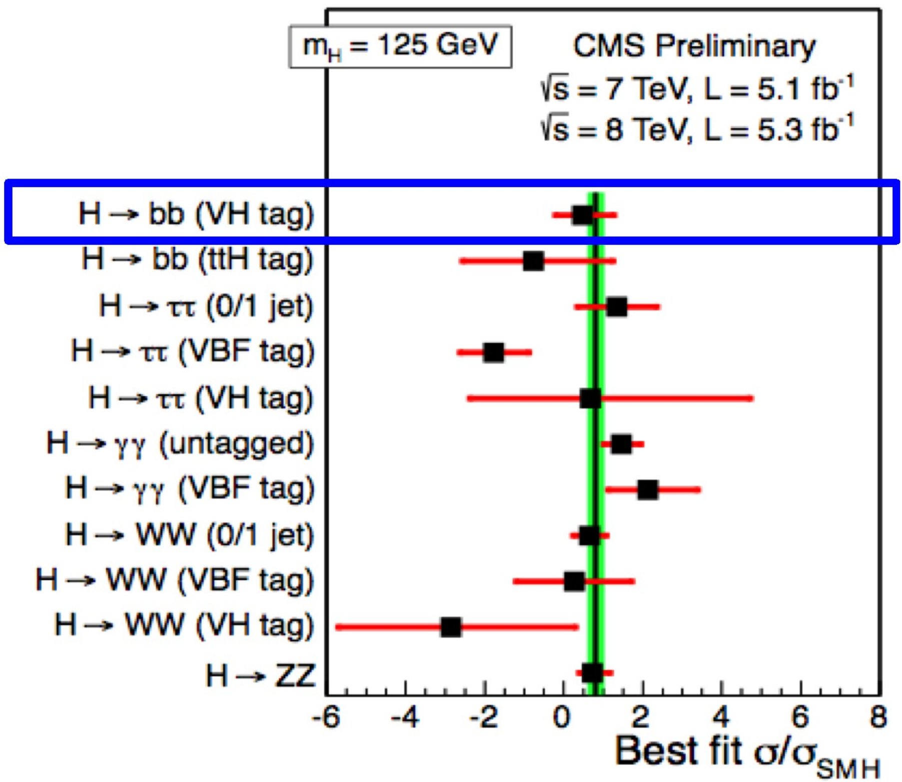


Signal injected at m_H=125 GeV

- Improvements in the analysis enhance sensitivity by 50%
  → Almost reached SM sensitivity (1.1 x σ_{SM}) below 115 GeV
  → Expected sensitivity around 1.6 x σ_{SM} for m_H=125 GeV
- Signal injected would give a broad excess across the full mass range considered
  → compatible with di-jet mass resolution (~10%)

Results: SM Exclusion Limits

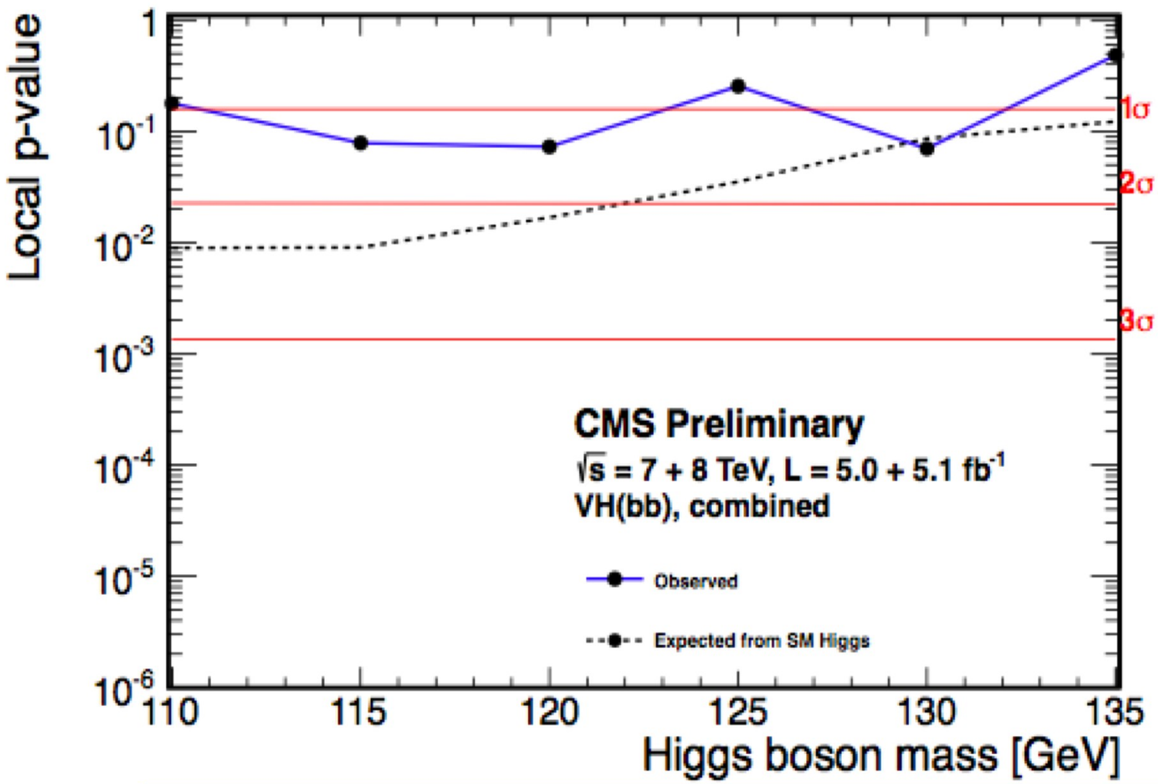


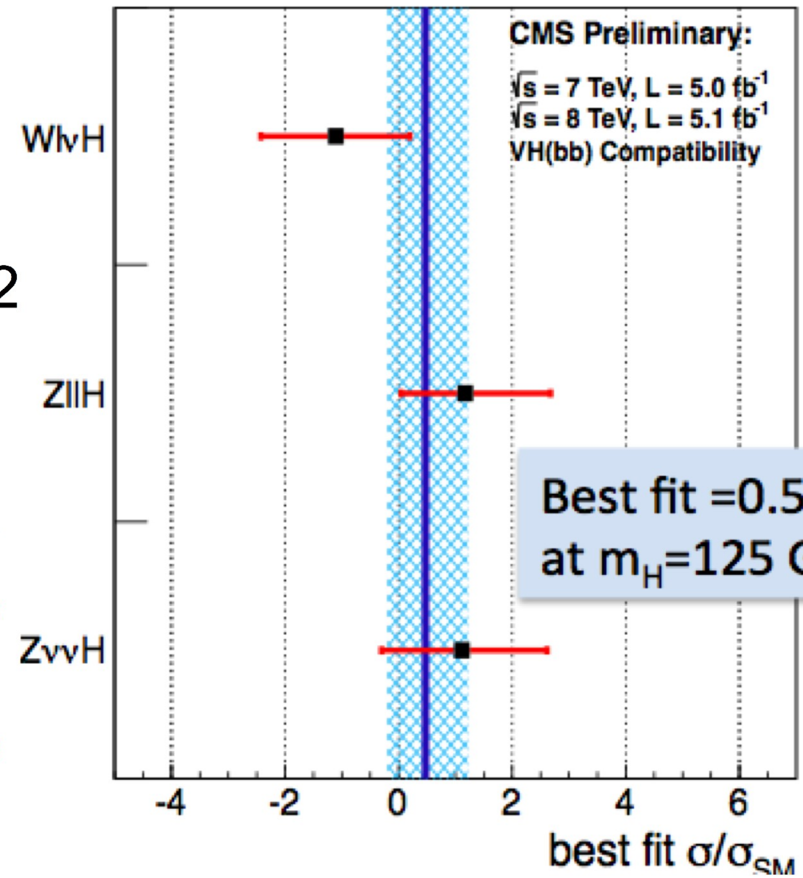


- Mild excess between 115 and 135 GeV
  → Most sensitive single experiment on VH → bb
- Compatible with either background or Higgs signal
  → Expect 1.6 x σ_{SM} at 125 GeV, observe 2. x σ_{SM}

Signal Strength and p-values

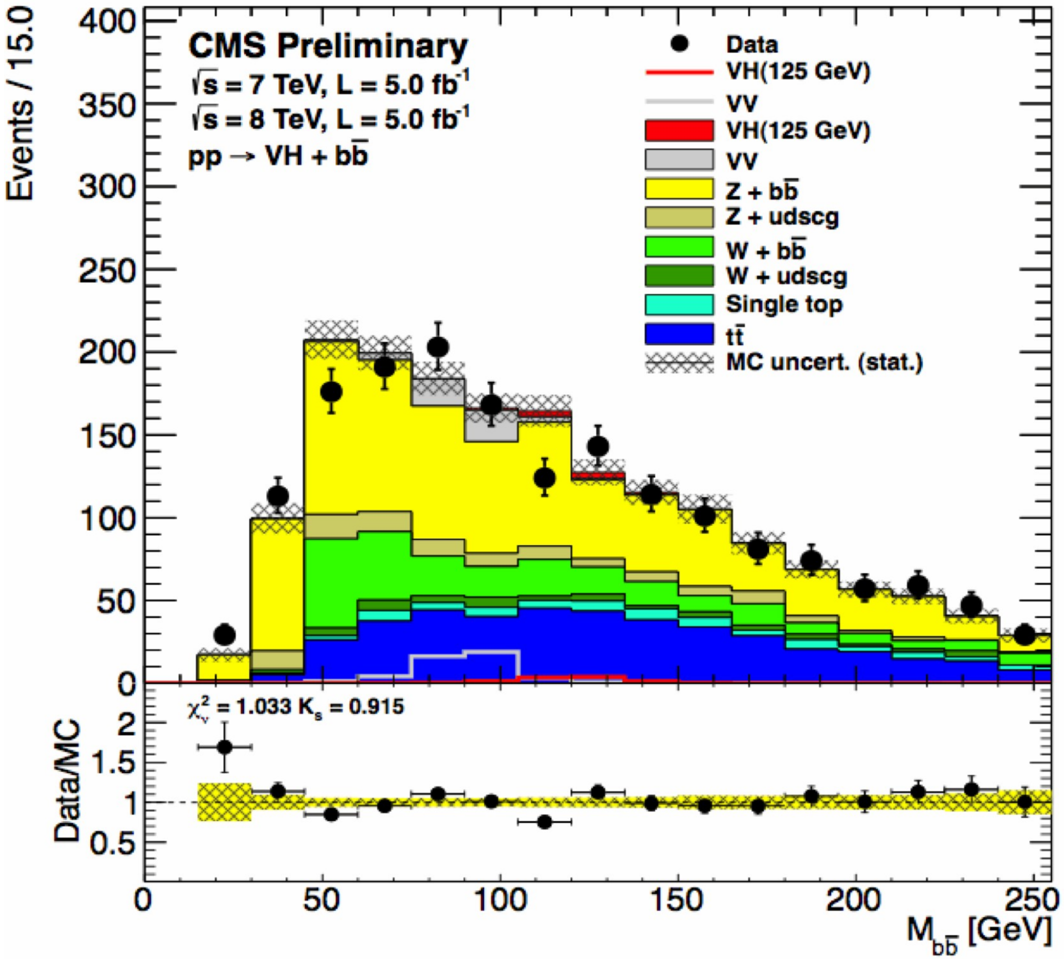
- Significance of the excess around 1σ in all mass range considered
- Looking forward to results on larger 2012 statistics!

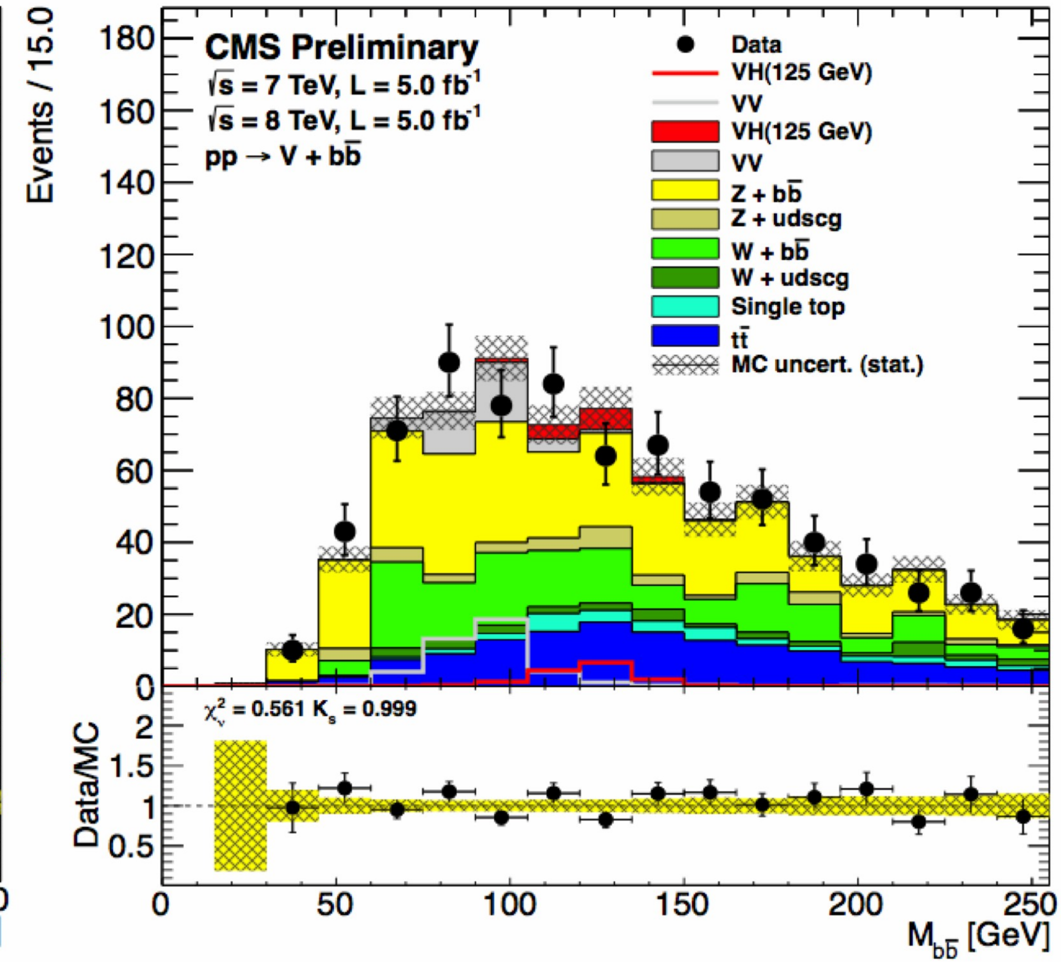




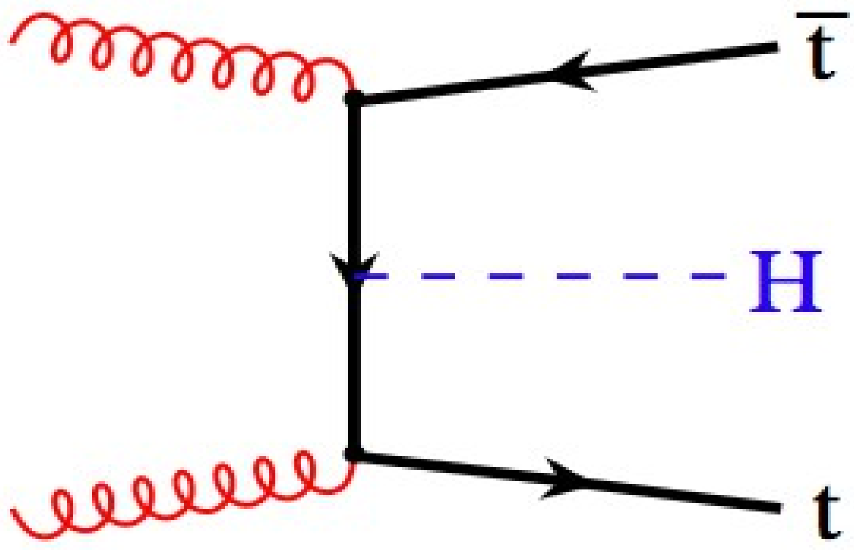
Best fit $=0.5\bullet\sigma_{SM}$ at m_H=125 GeV

7 + 8 TeV di-jet mass distributions





- Tighter cuts, stronger background rejection
- Show combination of 5 channels, overall nice Data/MC agreement



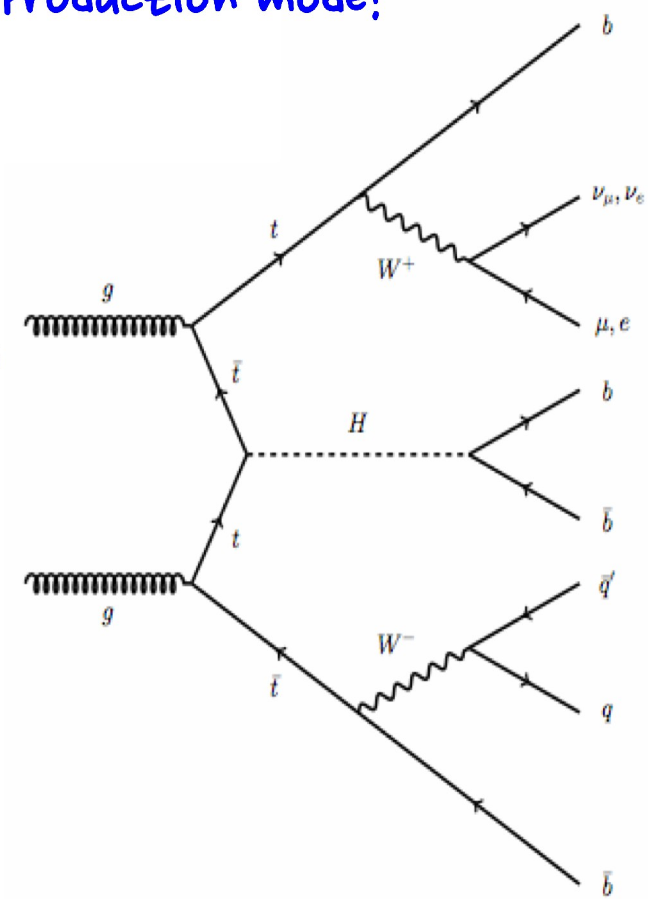
$$gg \rightarrow t\bar{t}H$$
$$H \rightarrow b\bar{b}$$

Presented for the first time at ICHEP
Test new production mechanism

ttH Analysis Overview

- Additional information in overall Higgs search
- Study lepton+jet (LJ) or di-lepton (DIL) top decays
- Major background from ttbar (+jet) events
- Split events by top decay and by number of jets and b-tags
- ANN to separate ttbar and ttbarH
  → Use simultaneous fit of ANN shape in each jet/tag category for search
  → Very different S/B, categories with low sensitivity help constraining B

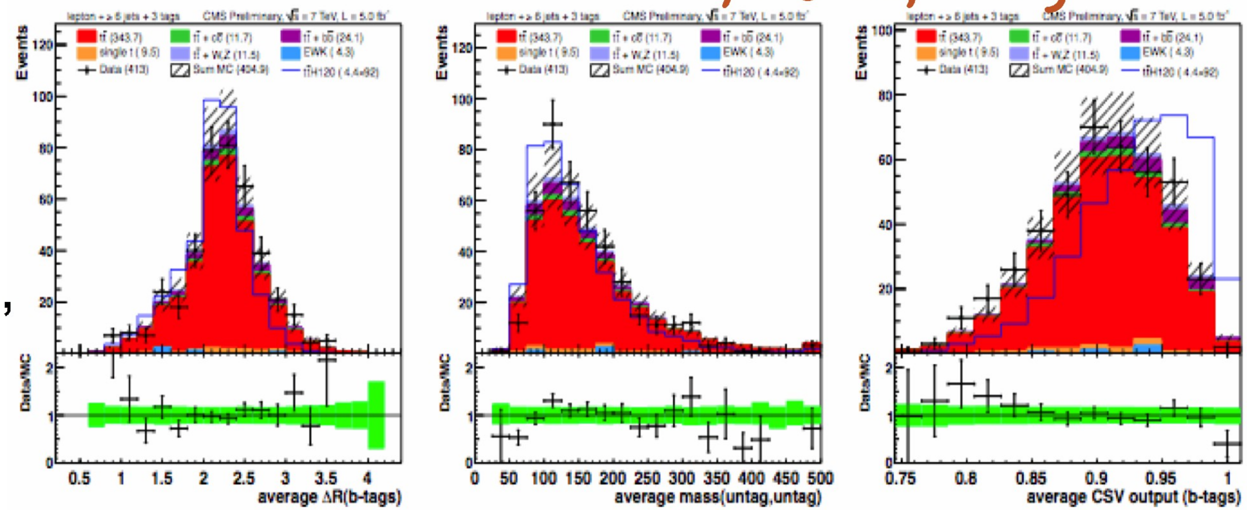
New Analysis,
First LHC study of this
Production mode!



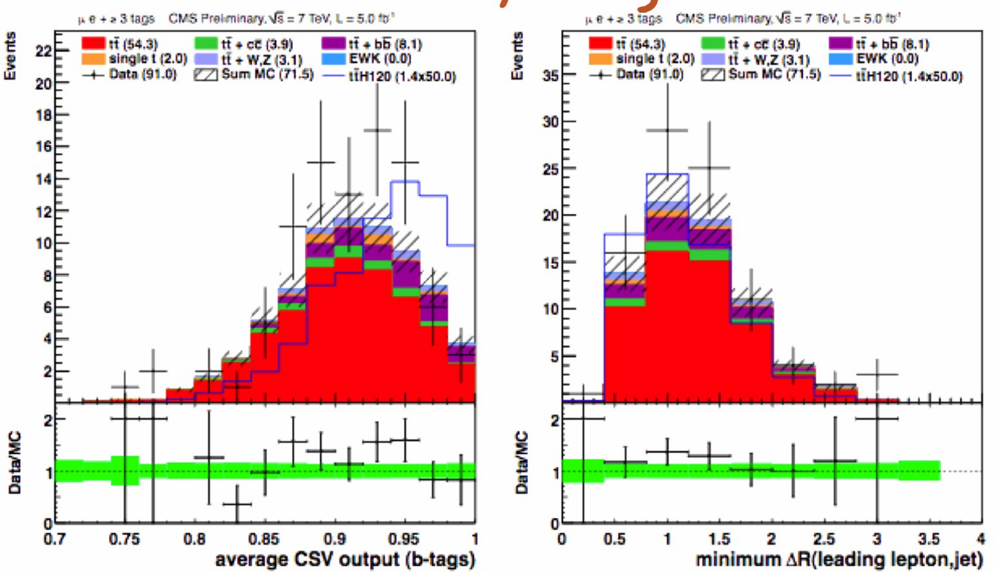
ANN Analysis Validation

- Build ANN discriminant for each (LJ or DIL) category
- Most relevant variables: b-tag, kinematic and angular correlation (e.g. min ΔR between all pairs of b-tagged jets)
  → Check data/MC agreement
- Irreducible background from tt+bb events studied with dedicated control region
  Built from ad-hoc ANN

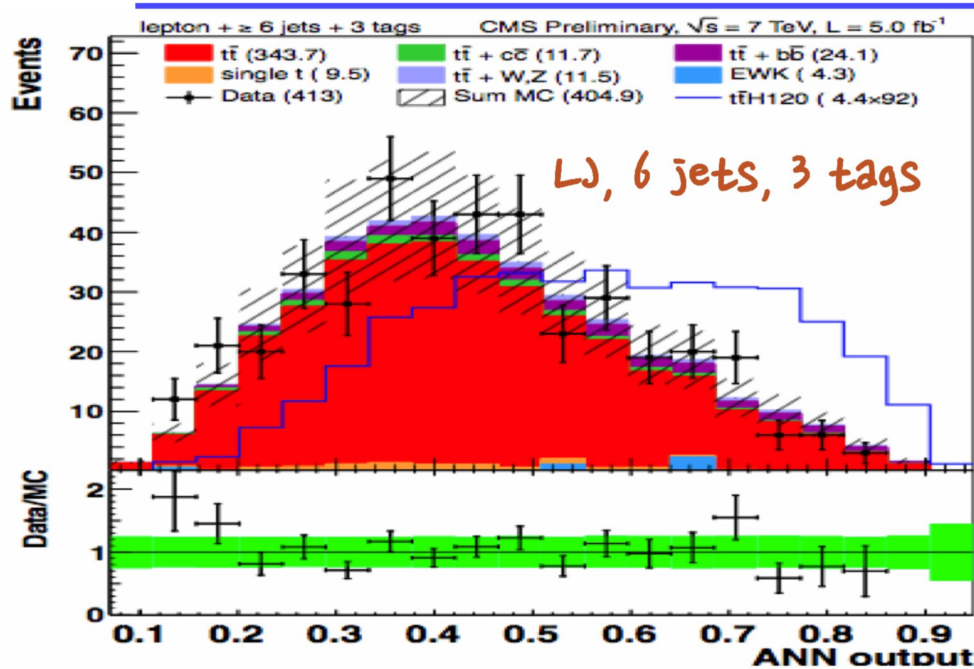
LJ, 6 jets, 3 tags

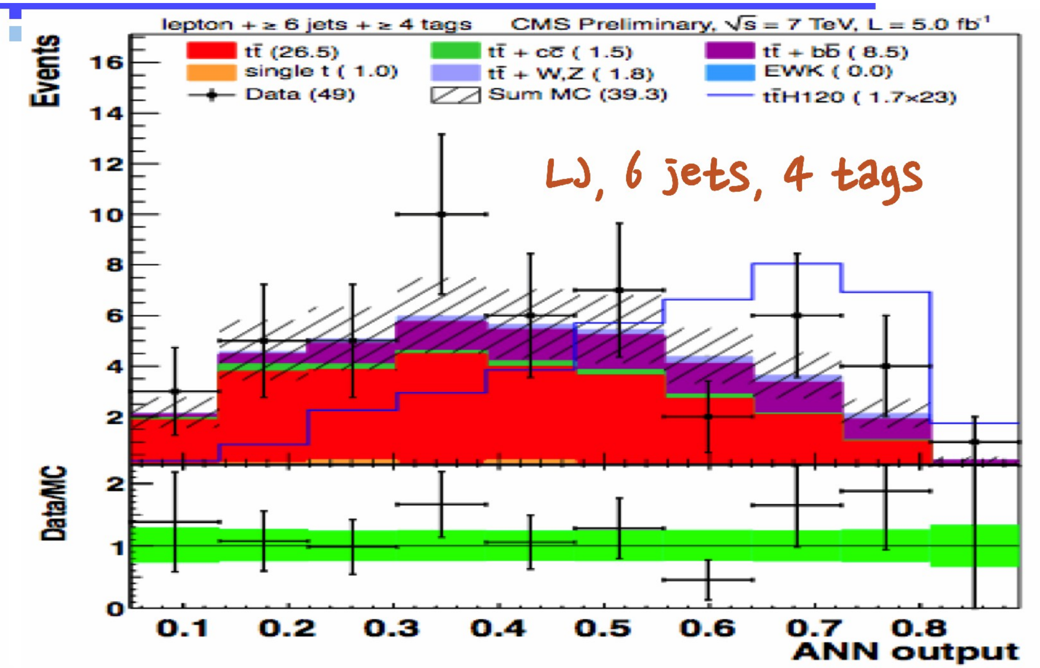


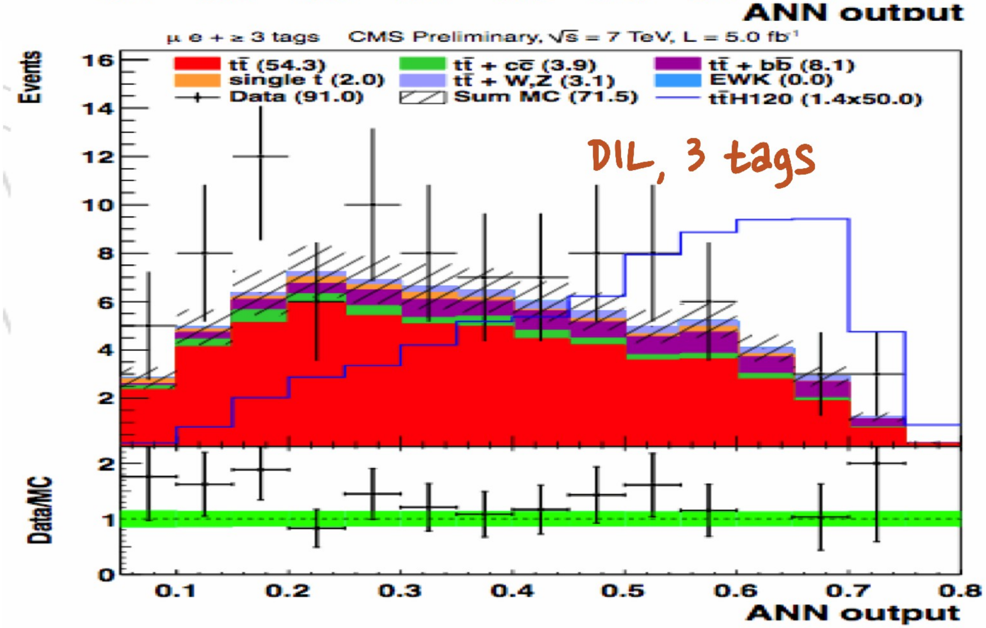
DIL, 3 tags



ANN Output Distributions



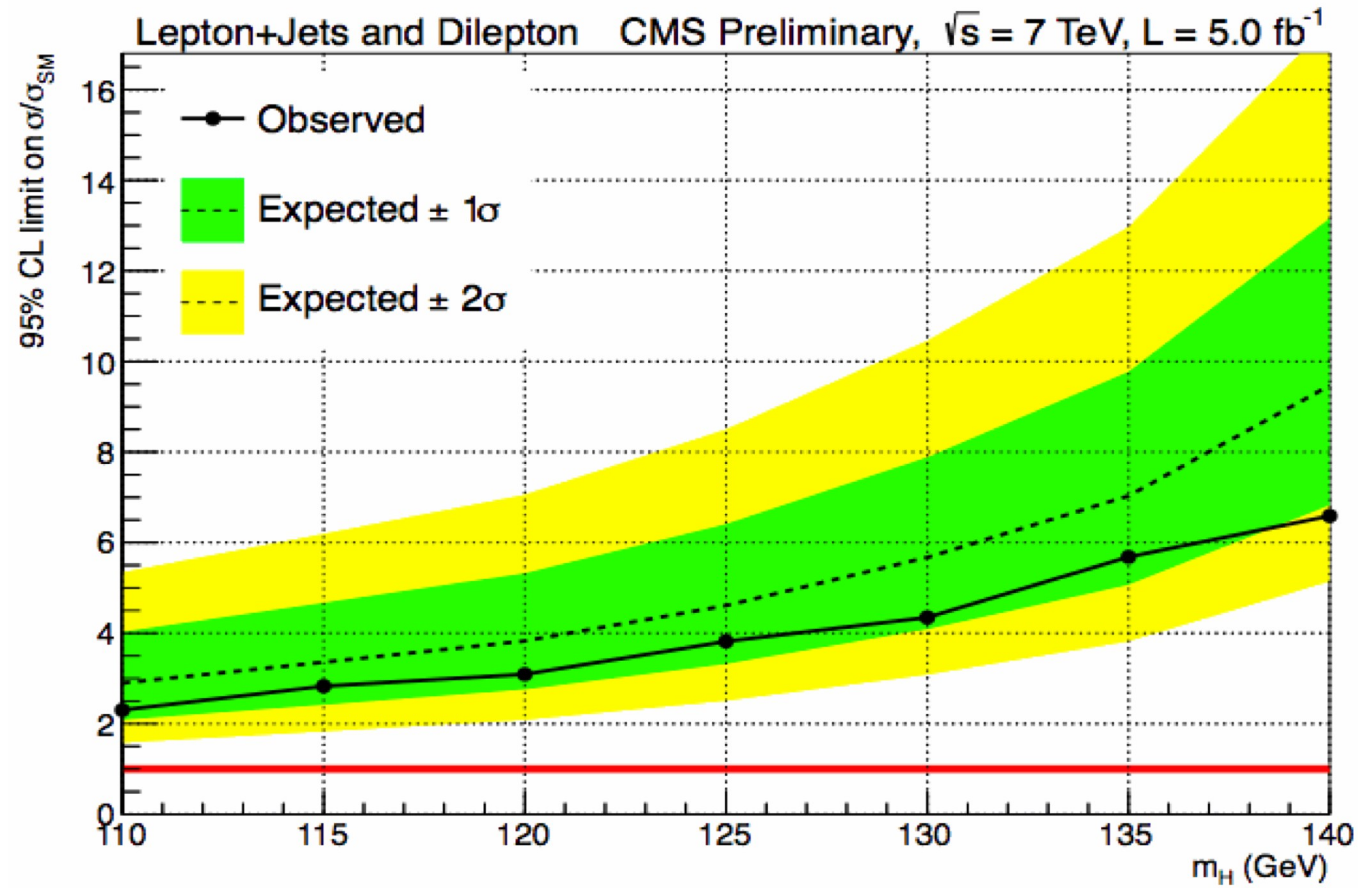




S/B strongly dependent on # tags
DIL: 2-3 tag categories
LJ: 2-4 tags, 4-6 jets

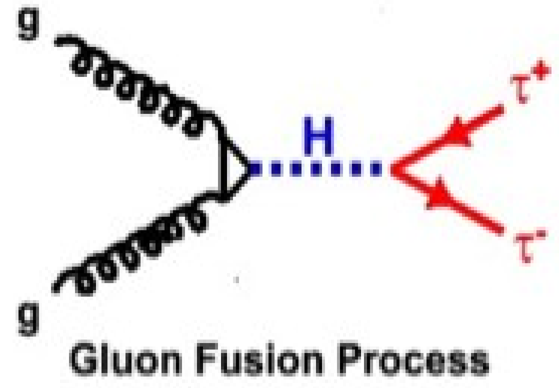
Signal expectation rescaled to
Σ (background)

Results: SM Exclusion Limits

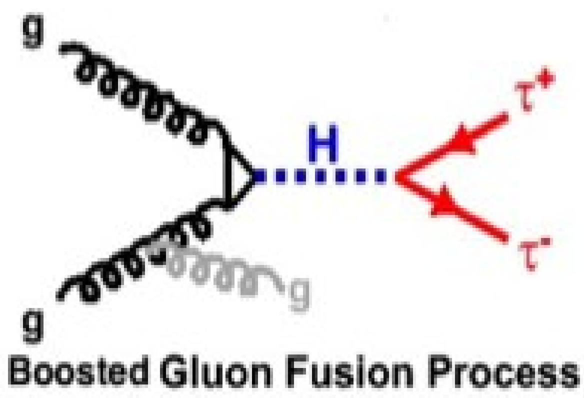


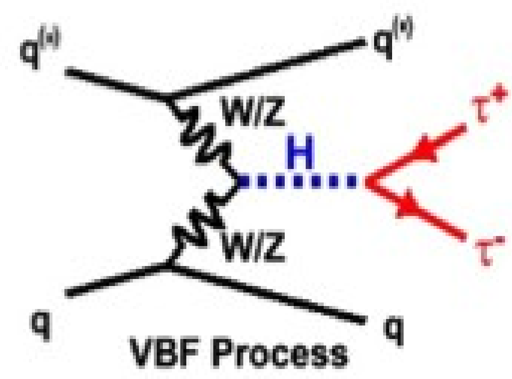
- Sensitivity dominated by lepton+jet mode, 5-10% improvement from di-lepton mode
- Dominant uncertainties: b-tag, JES in LJ, factorization scale in DIL
- No excess seen, expect 4.6 x σ_{SM} at 125 GeV, observe 3.8 x σ_{SM}

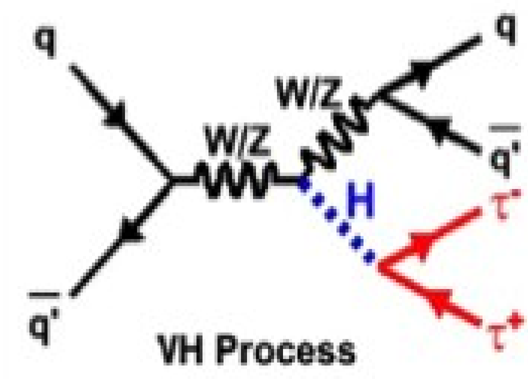
$H \rightarrow \tau\tau \rightarrow \mu\tau_h, e\tau_h, e\mu, \mu\mu$



Sensitive to all production modes
Probe couplings to leptons
Enhanced σ x BR in MSSM

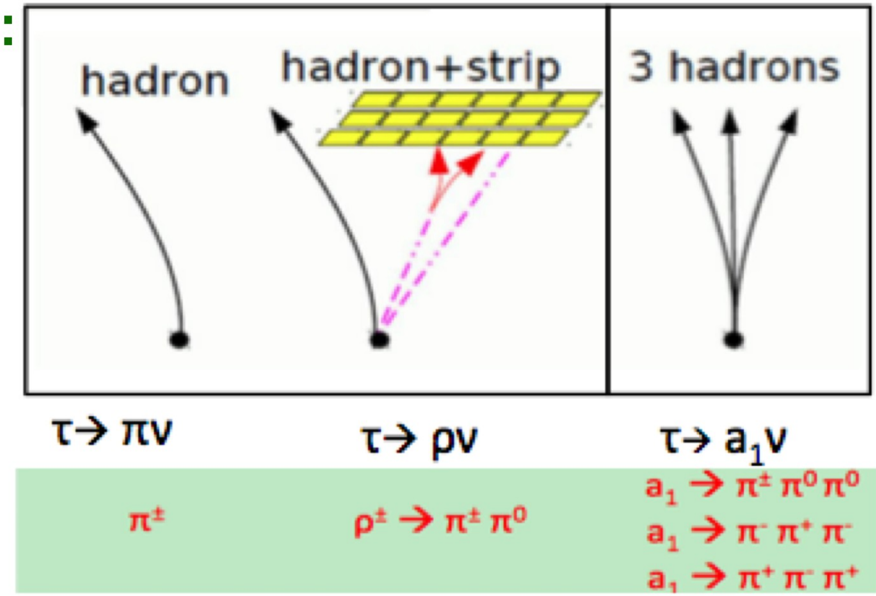


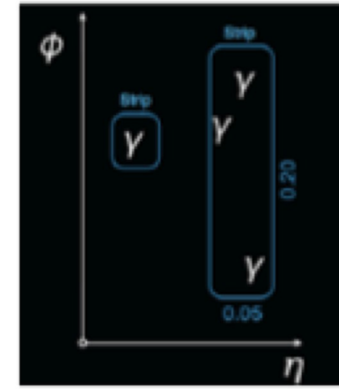




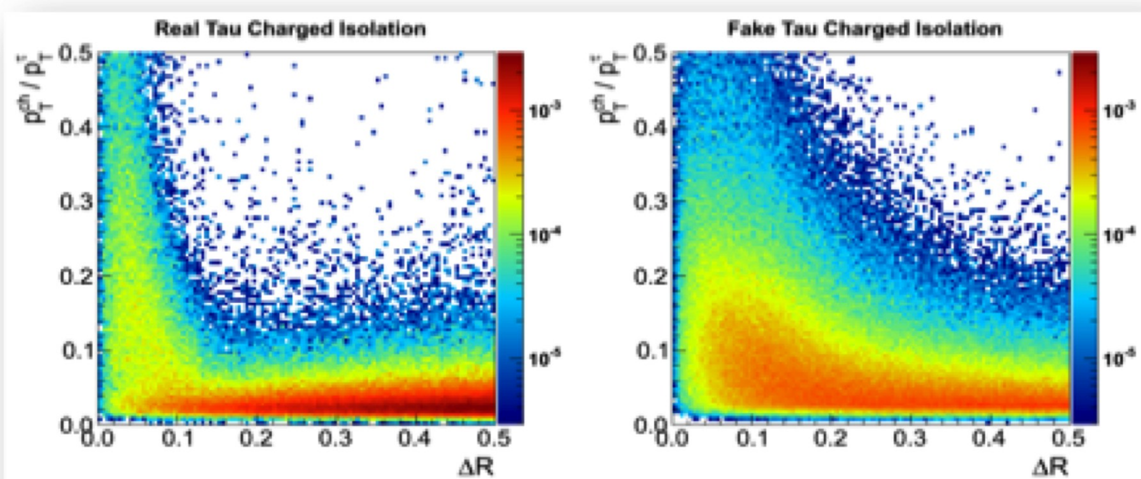
Taus at CMS

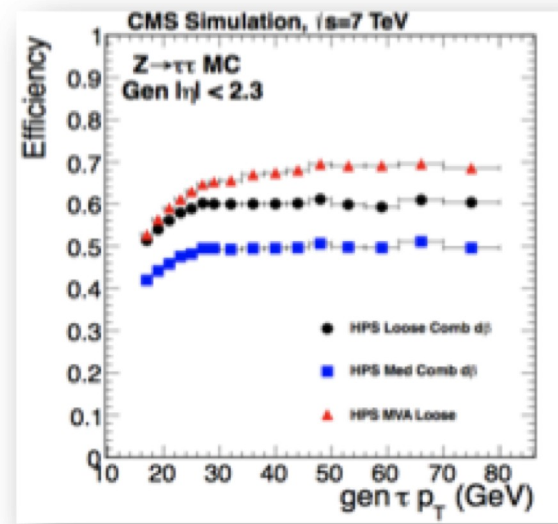
- **Hadronic taus identification:**
  Reconstruct individual decay modes:
  Charged hadrons + electromagnetic objects (arranged in strips or single photons)





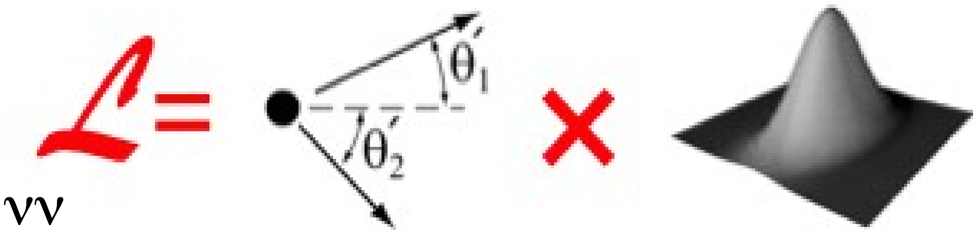
- **Tau isolation:**
  MVA discriminators using
  0.1<DR<0.5 annular deposits of energy



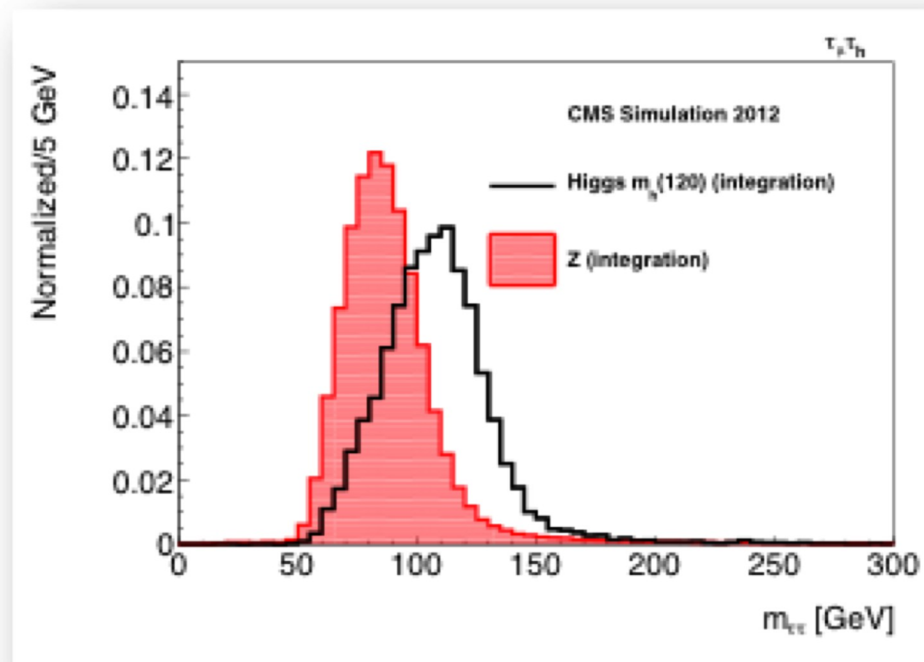


m(ττ) Reconstruction

- SVFit
  Event by Event estimator of
  true m(ττ) likelihood
  → Exact Matrix Element used for $\tau \to l\nu\nu$
  → Phase-Space is used for $\tau \to \pi$
  Nuisance parameters are integrated out

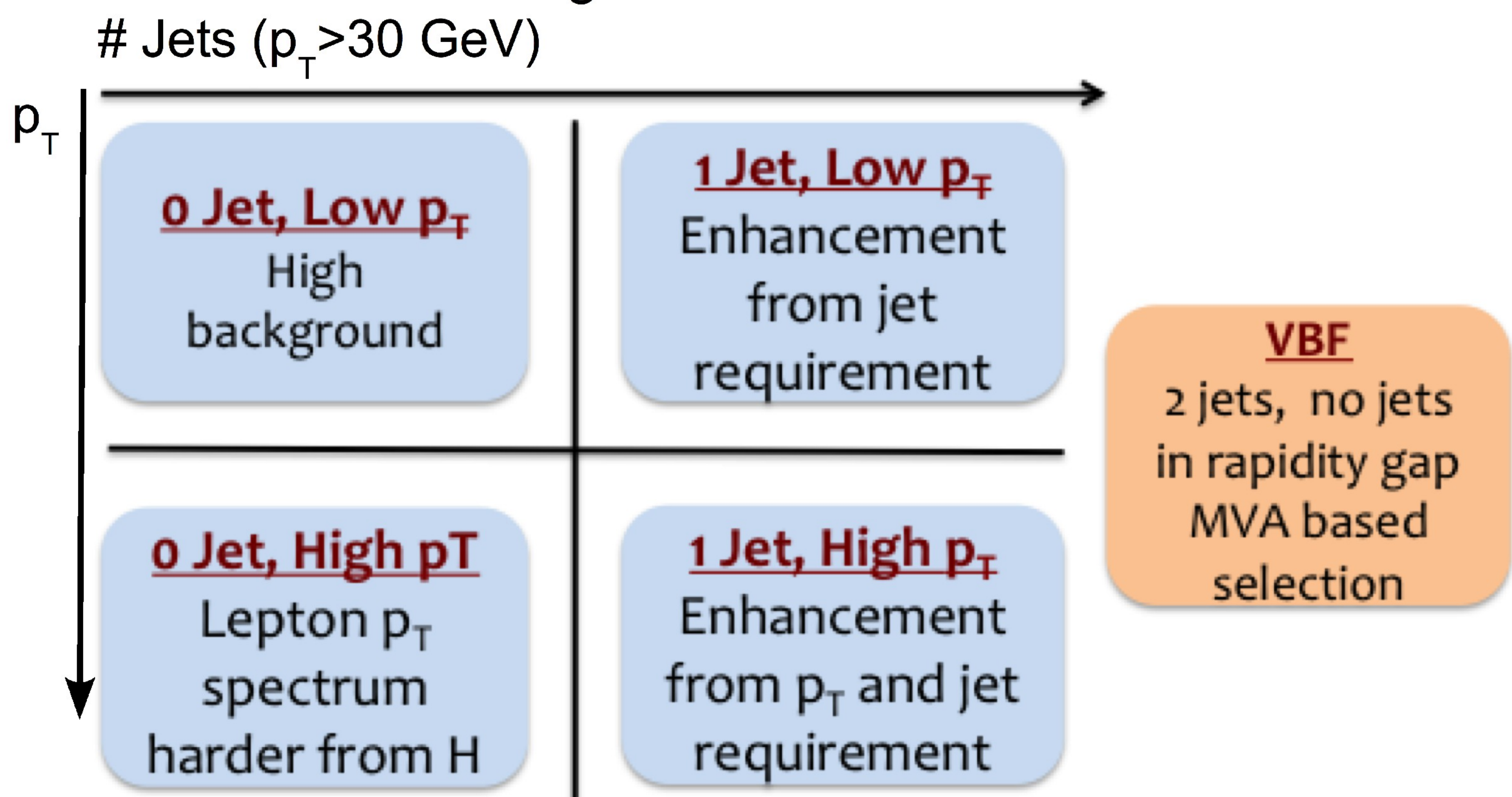


- Mass peaks at true value
  Mass resolution improved by 20%
  w.r.t. 2011 analysis
  Better separation between H/Z



Analysis Strategy

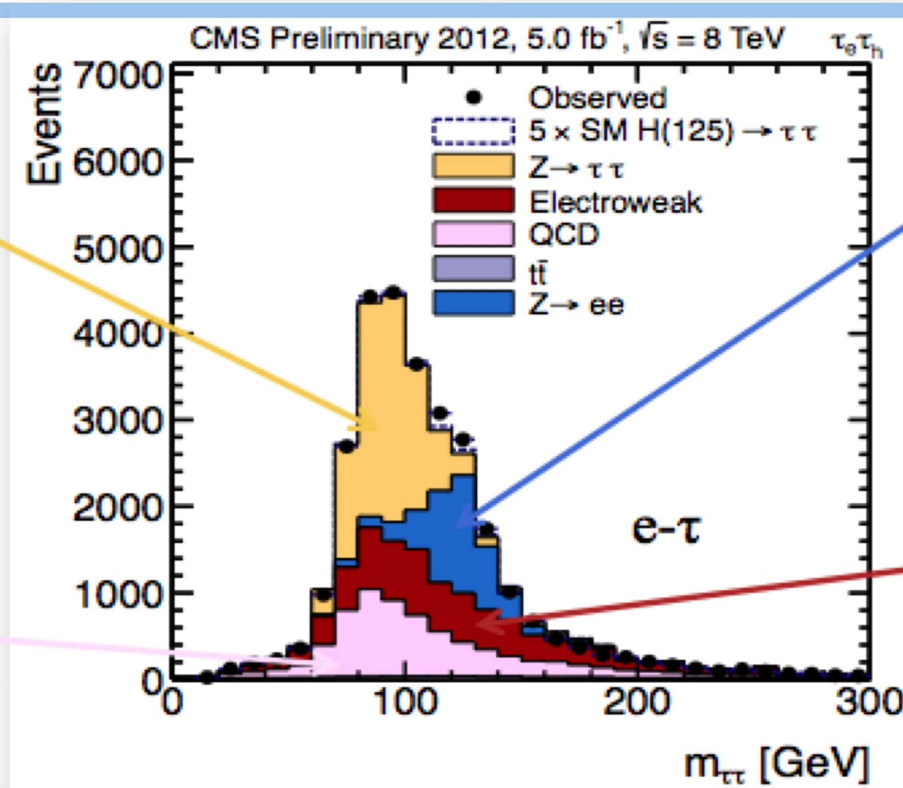
- Search performed in 4 tau-pair final states: $\mu\tau_h$, $e\tau_h$, $e\mu$, $\mu\mu$
- Analysis divided in 5 categories
  → Categorization based on $p_T(\tau_h)$ for , $p_T(\mu)$ for $e\mu$, leading $p_T(\mu)$ for $\mu\mu$
  → different S/B and mass resolution
  New since 2011 Analysis
- Simultaneous fit of all categories



Background Control

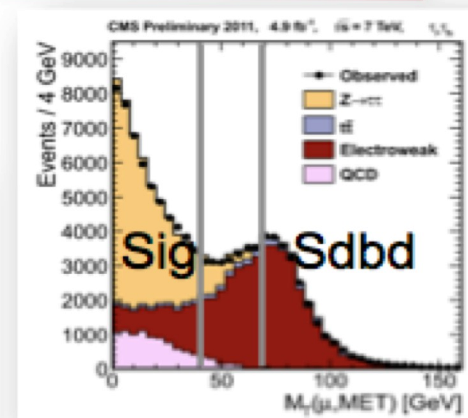
DY→ττ – Efficiency measured using τ embedded μμ events

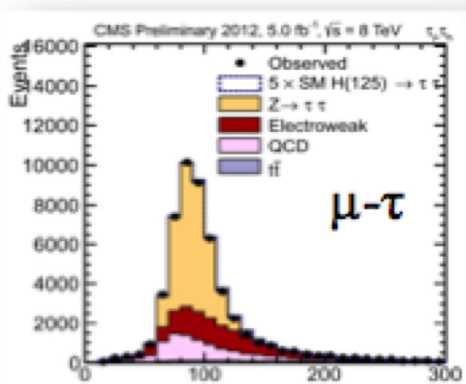
DY→ll – Taken from MC corrected for measured l→τ fake rates

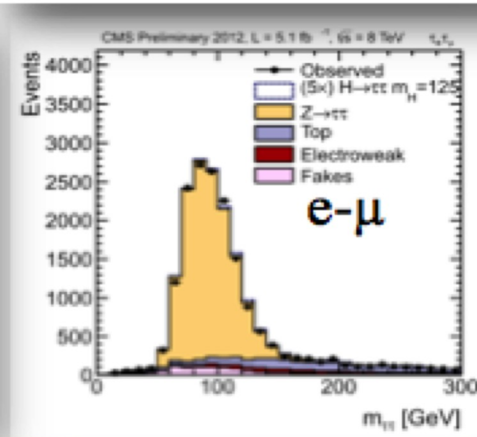


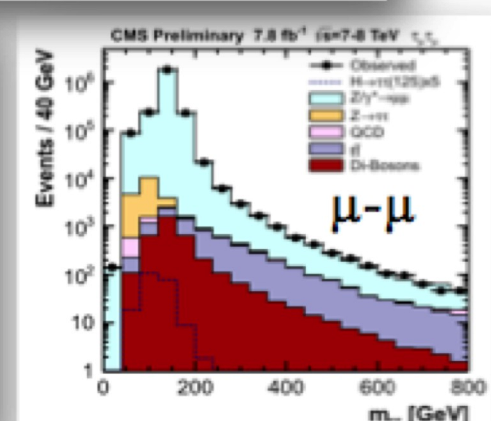
EWK – Mostly W+Jets, measured from high M_T sideband

QCD – Estimated from SS data





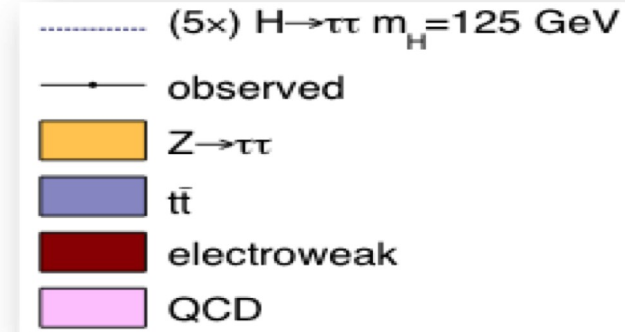


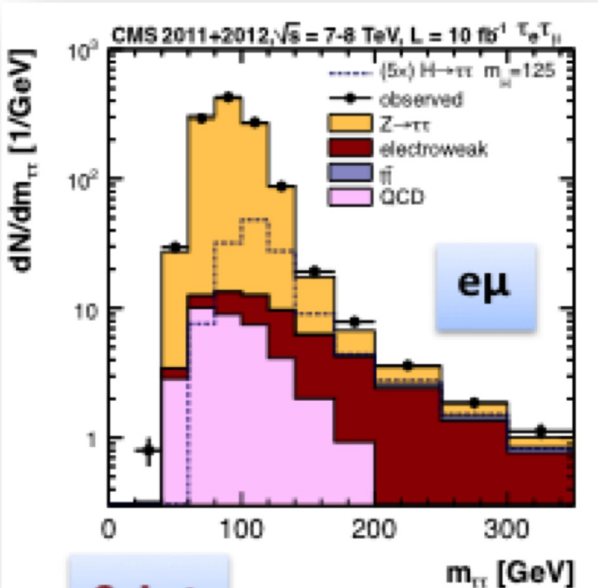


7/7/12

Plots are pre-fit

$m(\tau\tau)$ Categories

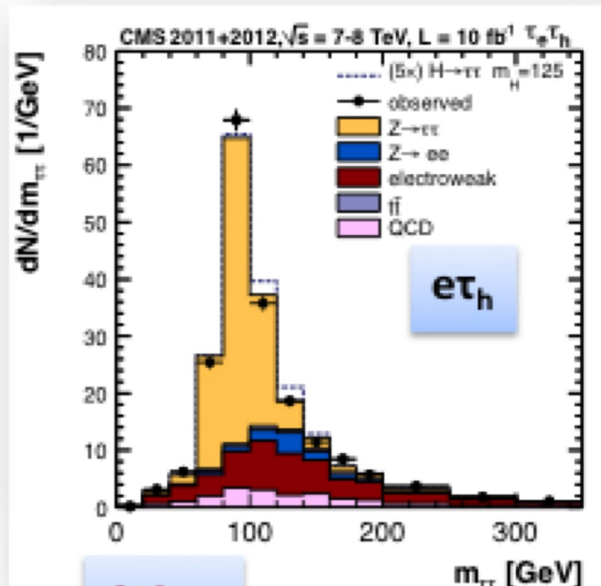




0 Jet

Bulk of events fall in this category

→ Sensitivity boosted by low/high p_T split

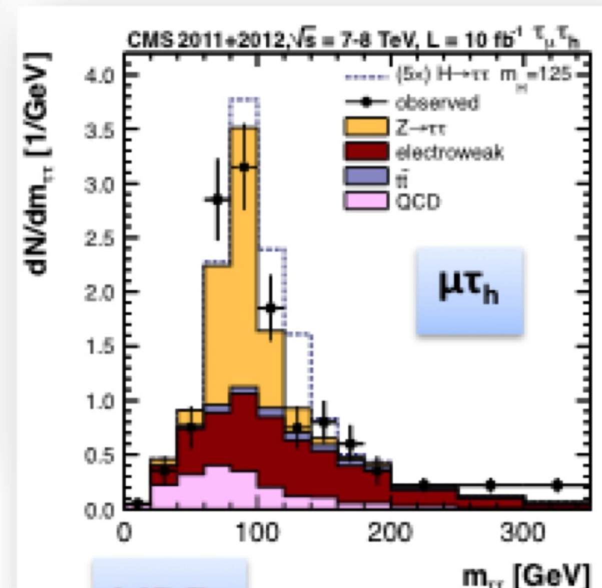


1 Jet

Enhances gluon- gluon fusion production

→ Improves mass resolution

→ High/Low p_T split makes this a powerful category

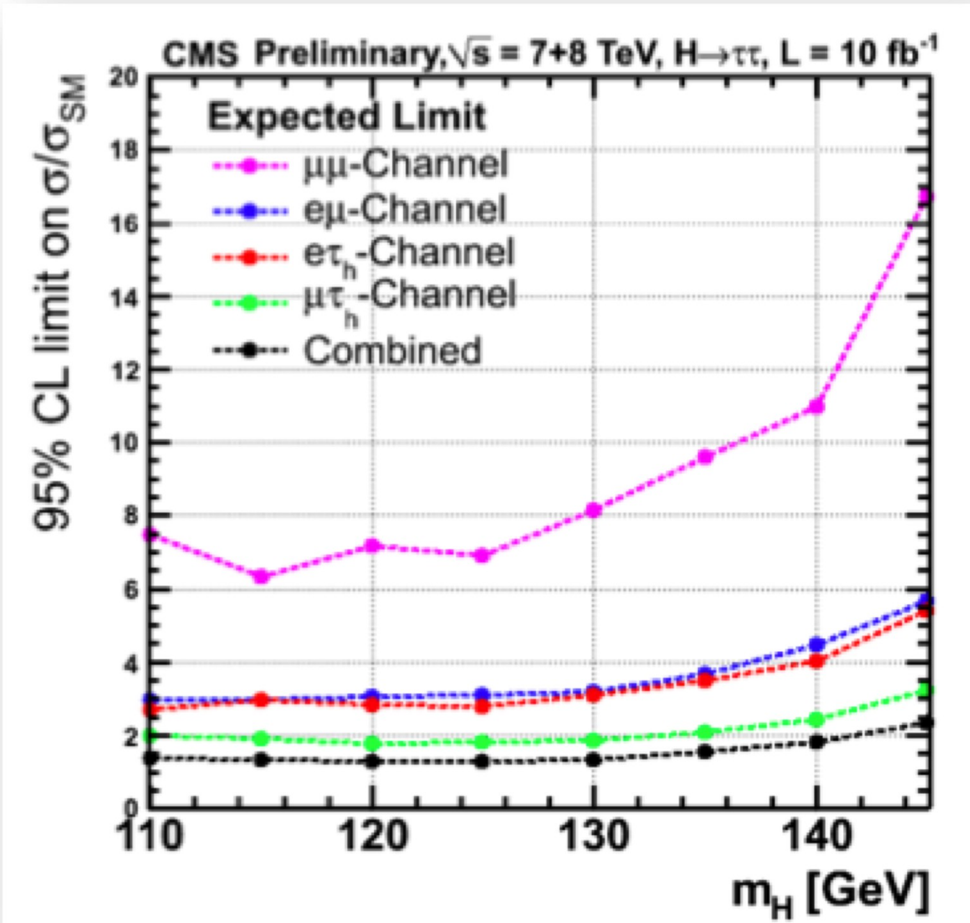


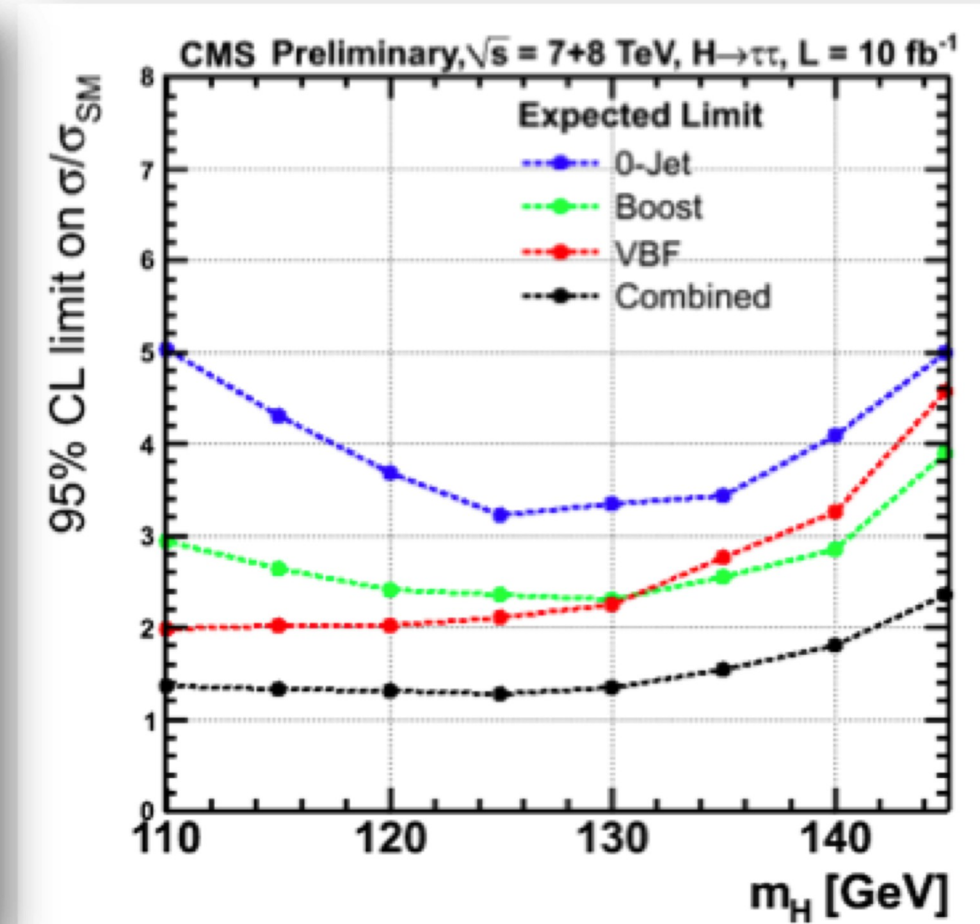
VBF

Enhancement for VBF production

→ Highest sensitivity channel for M_H < 130 GeV

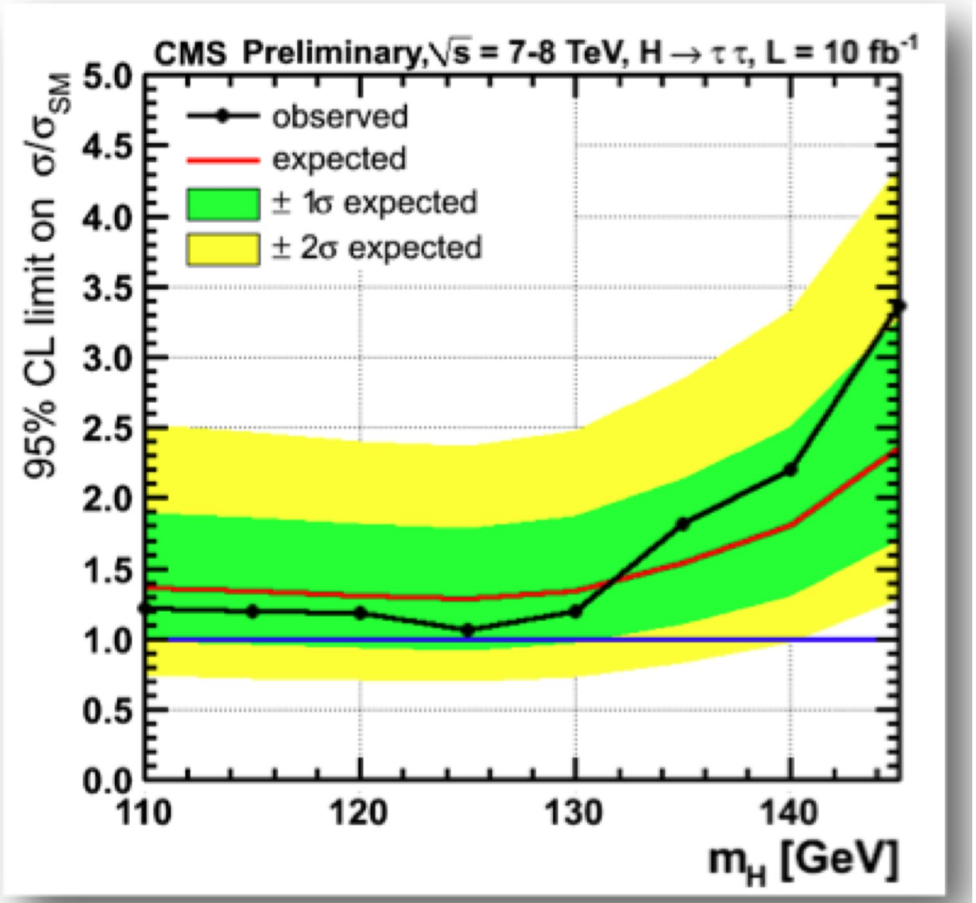
Summary of expected Limits

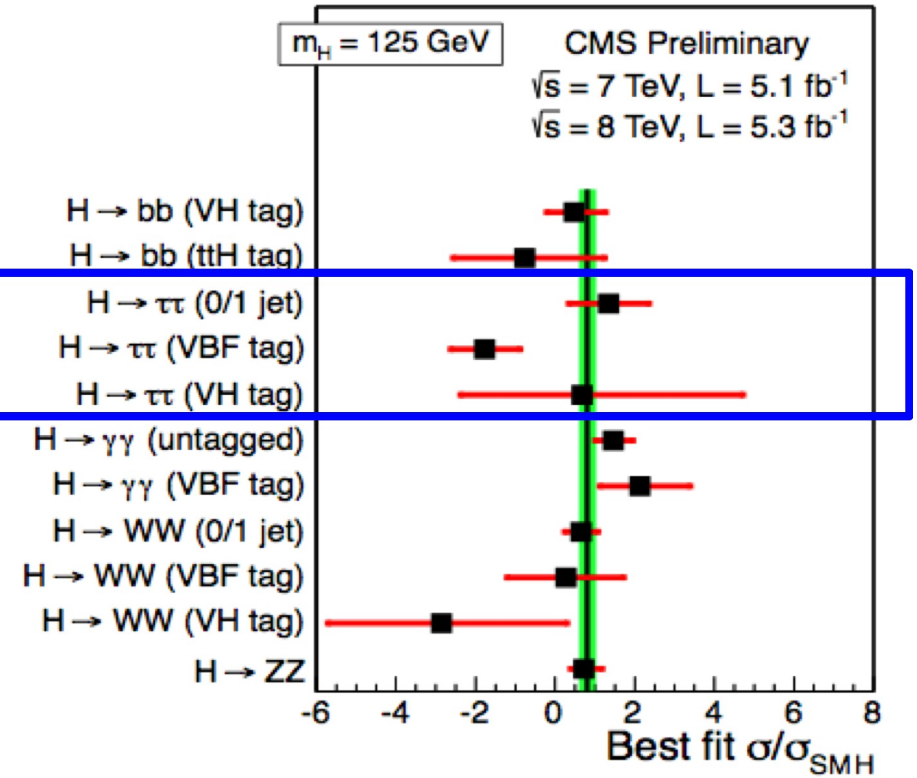




- **Analysis improvements make ττ a potent Higgs search channel**
  → Improved Categories, Mass Resolution
  → 2x improvement from 2011 published analysis

$H \rightarrow \tau\tau$ Results

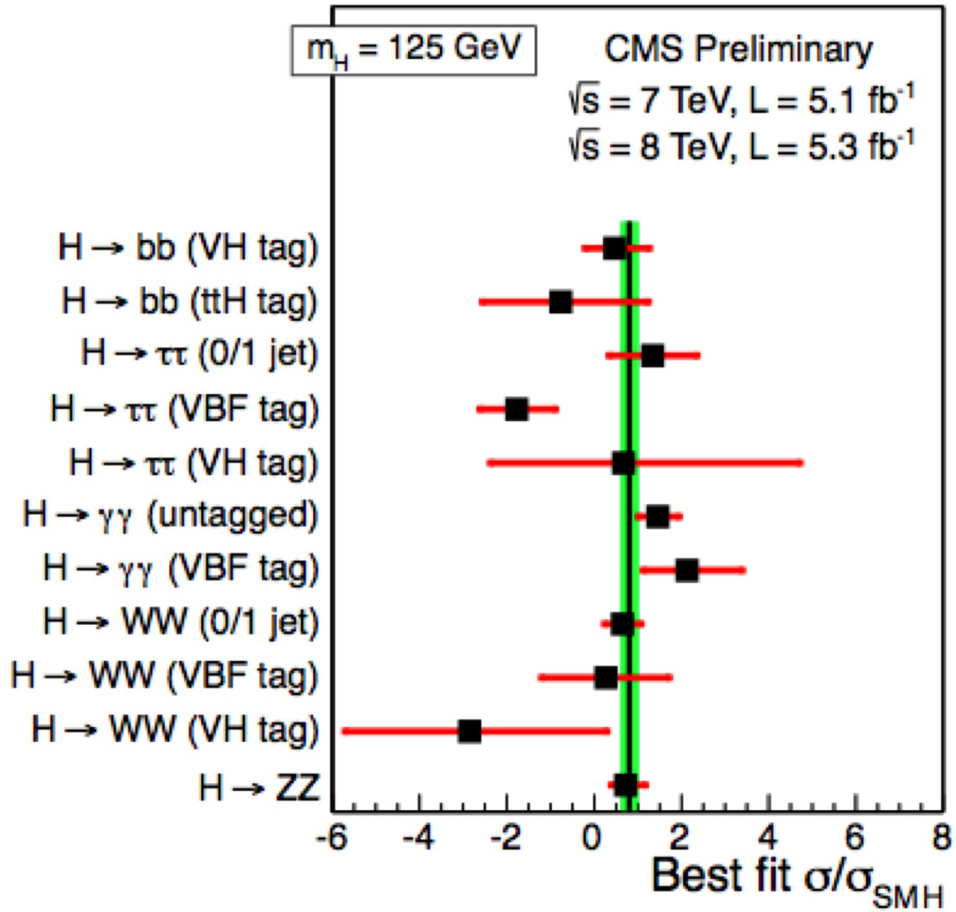




- Sensitivity close to Standard Model one!
- No excess seen from SM background-only expectation
- Observed limit of 1.06 x σ(SM) at m_H =125 GeV (exp= 1.3)

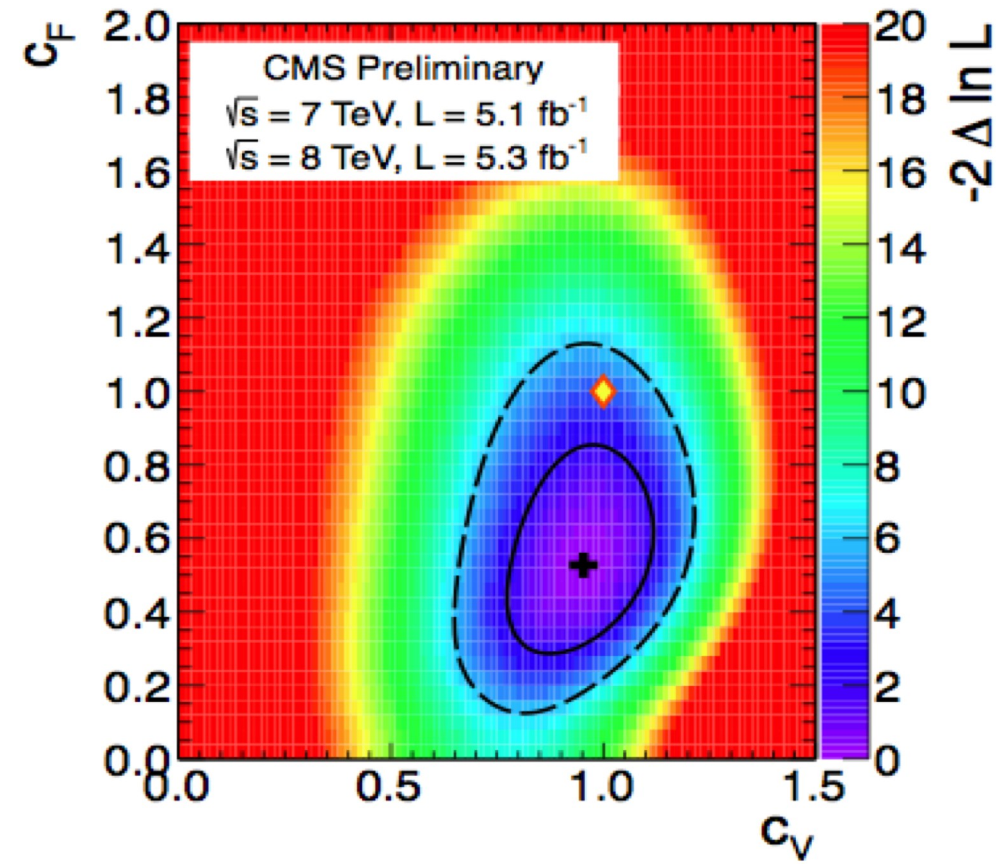
→ Under-fluctuation in VBF category drives the observed limit

The Global Picture



- Measuring coupling separately in vectorial and fermionic modes (C_F,C_V)
- Assume SM higgs @125 GeV

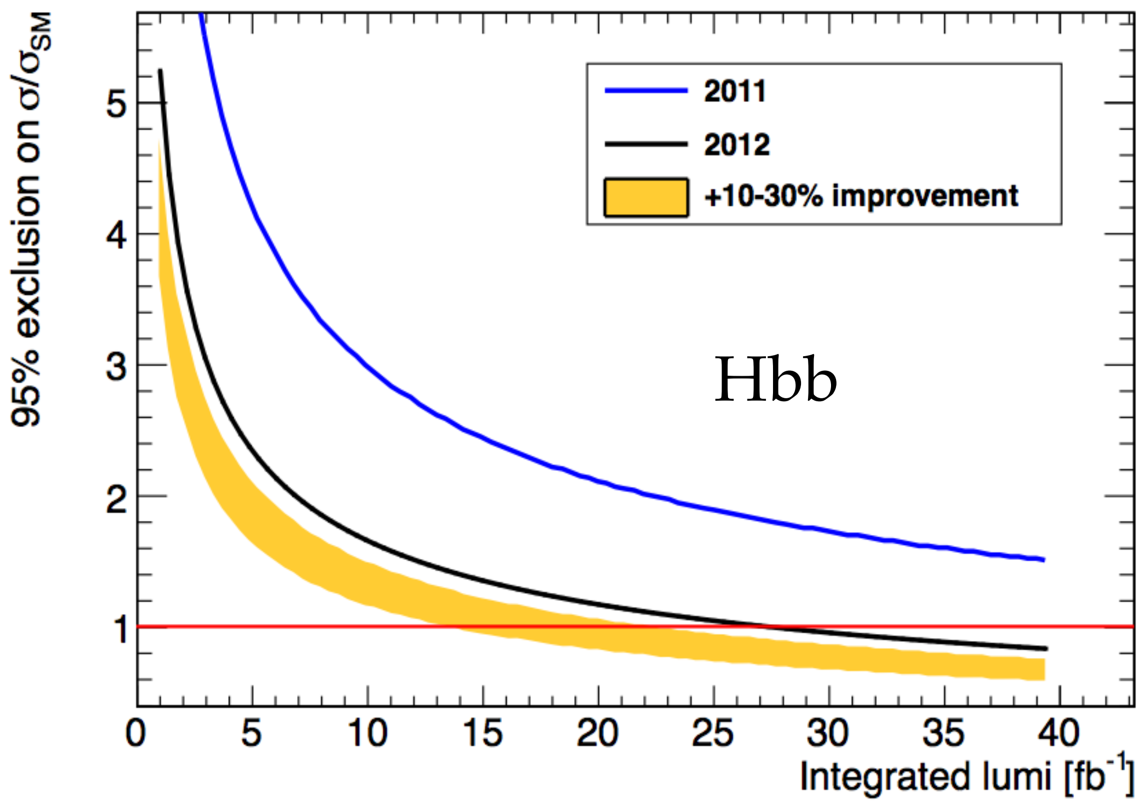
Results driven by under-fluctuation in ττ

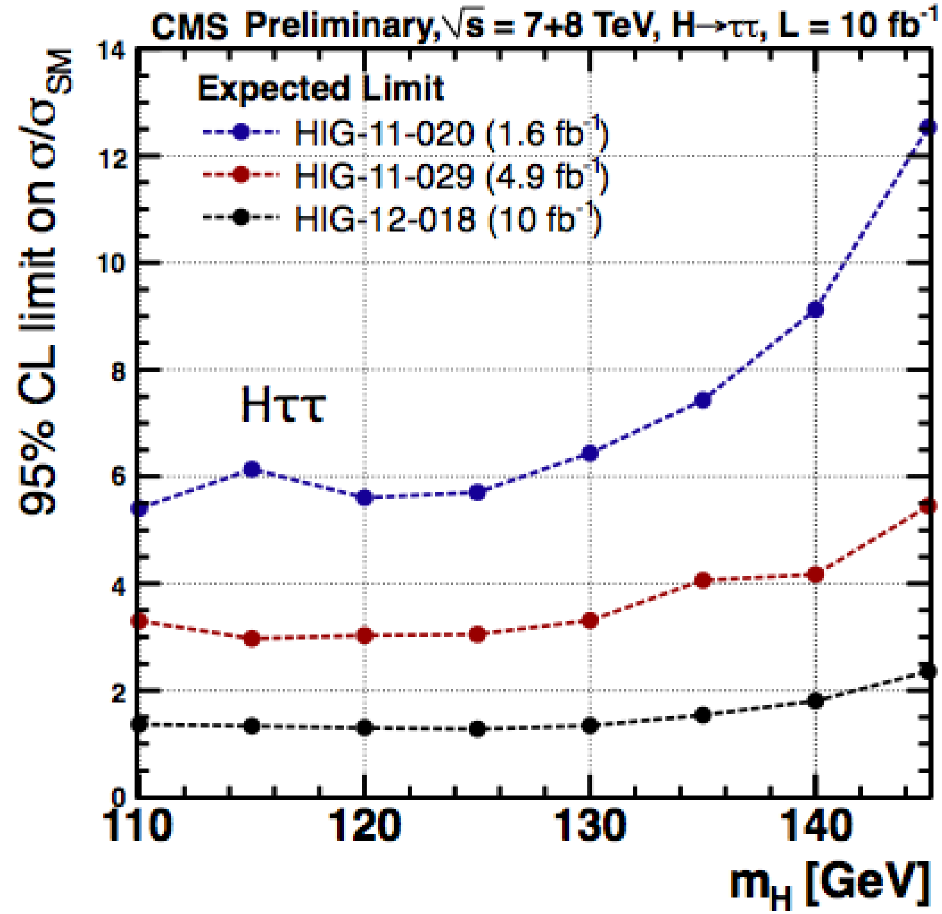


- Best fit consistent with SM

→ Excess in γγ compensated by some "outliers"

Prospects





- Analyses improve better than √ L
- Already ≤ 15 fb^{-1} at √s 8 TeV on tape
- Also large room for improvement in ttH

Conclusions

- The characterization of the new boson discovered at m=125 GeV at the LHC is top the priorities of the CMS and ATLAS physics programs
- Outstanding performances of the LHC should allow to shed some light on the nature of this new particle by the end of the year
- Presented most recent results on search for SM $H \rightarrow b\bar{b}$ and $H \rightarrow \tau\tau$ in CMS
  $\rightarrow$ Test coupling to fermions
  $\rightarrow$ $H \rightarrow b\bar{b}$ largest BR for m_H=125 GeV
- Mild excess in $H \rightarrow b\bar{b}$, under-fluctuation in $H \rightarrow \tau\tau$
  Stay tuned for updates in these channels!

Backup Slides

Data Samples and Triggers

- Analysis presented here based on full 2011 data sample (5 $fb^{-1,}$ VH+ttH) and 2012 Data collected until June TS (5 $fb^{-1,}$, VH)

2011

Mode	Lepton Trigger	Cross-Trigger (Jet, MET)
$W(\mu\nu)H$	(Isolated) muon, 17-40 GeV	-
$Z(\mu\mu)H$	(Isolated) muon, 17-40 GeV	-
$W(e\nu)H$	Isolated electron, ID cuts, 17-32 GeV	2 jets (25-30 GeV) + MHT (15-25 GeV)
$Z(ee)H$	Di-electron, 17-8 GeV	-
$Z(\nu\bar{\nu})H$	-	MET (80-100 GeV) + 2 jets (20 GeV) OR MHT (150 GeV)
$t\bar{t}H$	Isolated muon, 24 GeV	-
$t\bar{t}H$	Isolated electron, ID cuts, 25 GeV	3 jets (30 GeV)
$t\bar{t}H$	two leptons (electron and/or muon), 17-8 GeV	-

2012

Mode	Lepton Trigger	Cross-Trigger (Jet, MET)
$W(\mu\nu)H$	(Isolated) muon, 24-40 GeV	-
$Z(\mu\mu)H$	(Isolated) muon, 24-40 GeV	-
$W(e\nu)H$	Isolated electron, ID cuts, 27 GeV	-
$Z(ee)H$	Di-electron, 17-8 GeV	-
$Z(\nu\bar{\nu})H$	-	MET (80 GeV) + 2 jets (25-60 GeV), $\Delta\phi$ cuts OR MHT (150 GeV)

- Lepton efficiencies determined directly on data using Z events
InVH, trigger Efficiencies well above 90% w.r.t. offline cuts (Boost)

Data Samples and Triggers

- **Analyses presented here:**
  Associated production with a vector boson (VH, V=W,Z):
  Improved Analysis of 2011 data (5 fb^{-1}) and first analysis of
  2012 Data at $\sqrt{s}$=8 TeV (5 fb^{-1})

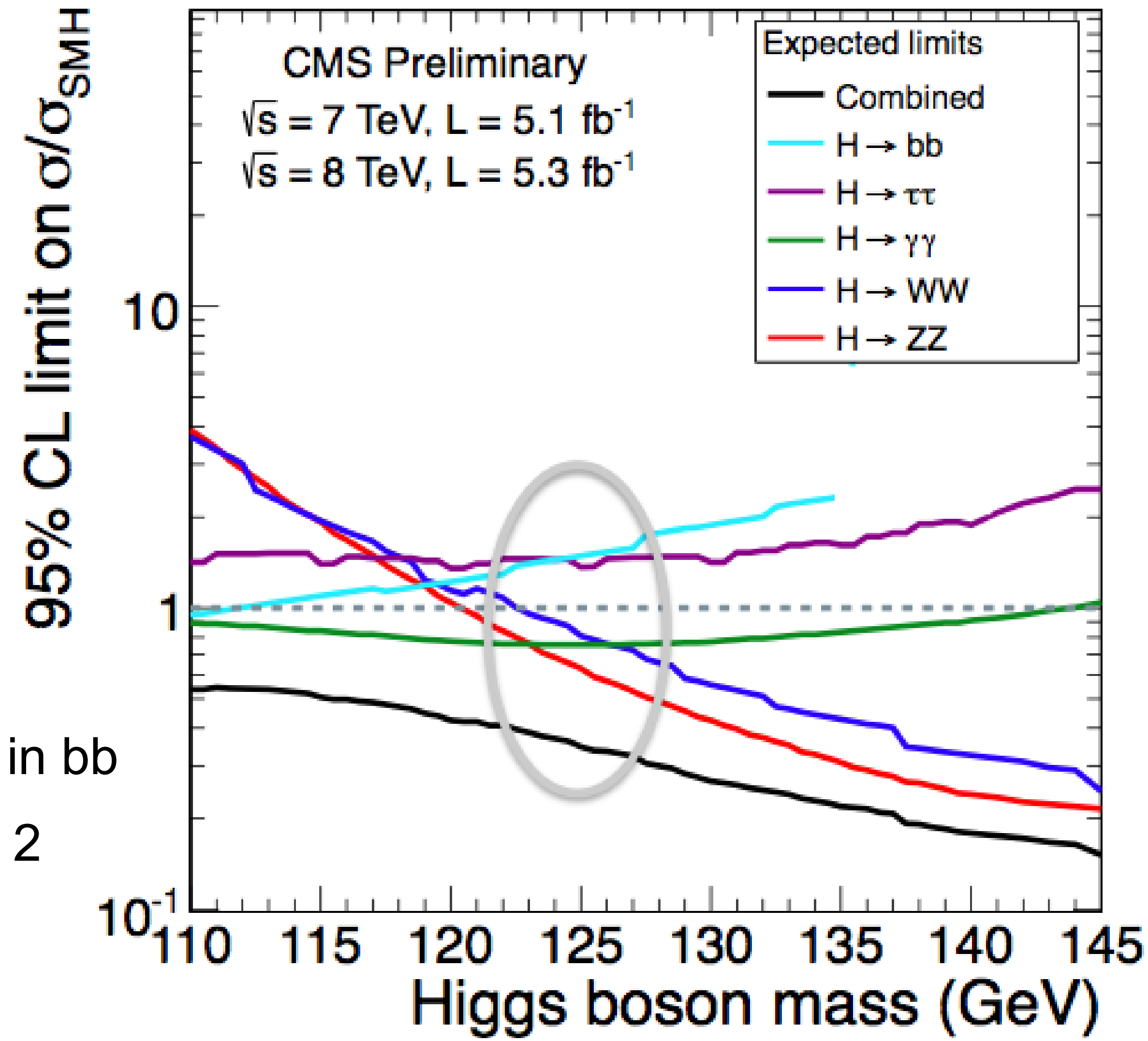
- **Triggers:**
  (Isolated) muon, 17-40 GeV (2011), 24-40 GeV (2012) → $W(\mu\nu)H, Z(\mu\mu)H$

  Isolated electron, 17-32 GeV (2011), 27 GeV (2012) → $W(e\nu)H$, ttH
  → Cross-trigger with central jets and MET in 2011

  Double lepton, 17-8 GeV → Z(ee)H, ttH

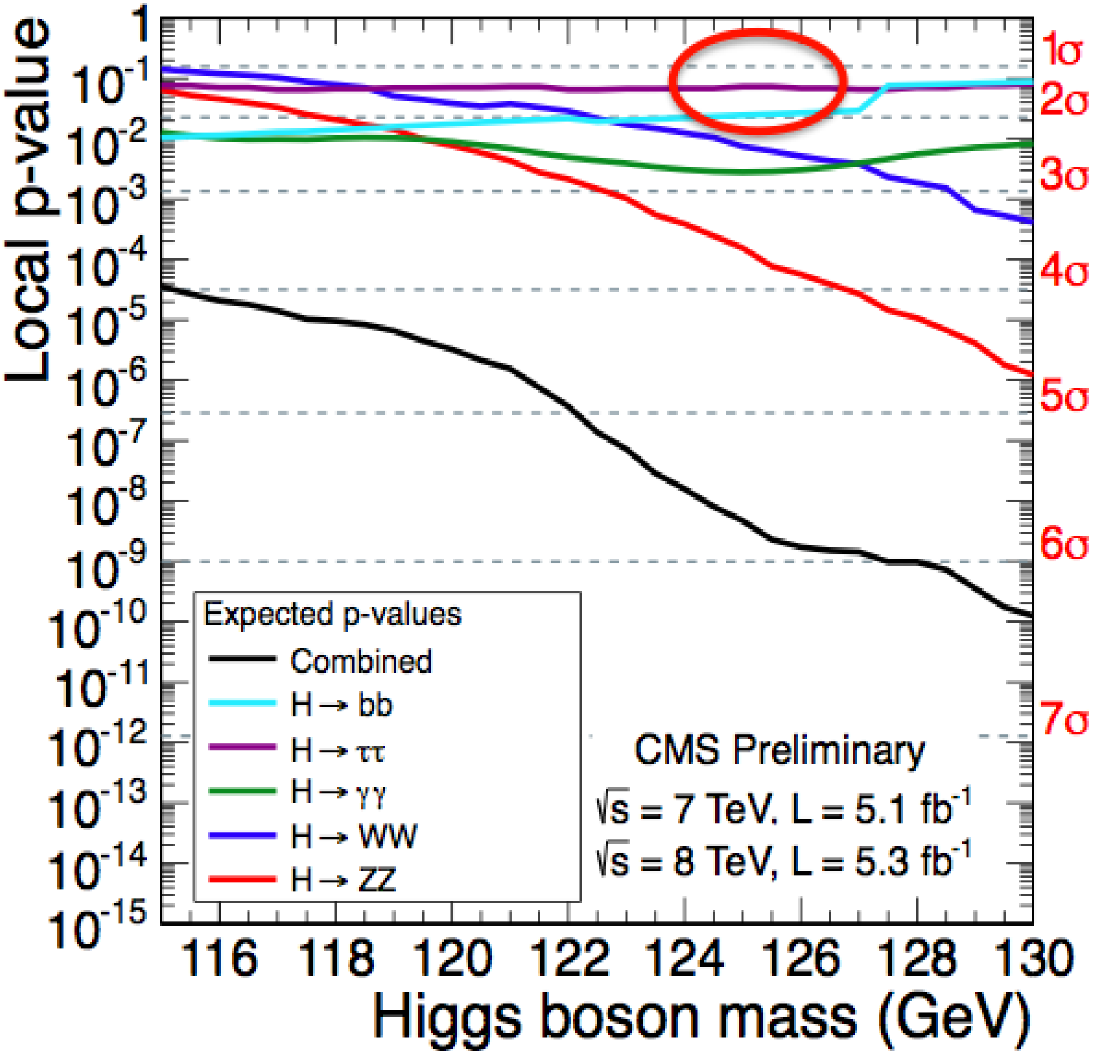
  MET (80-100 GeV) with central jets or inclusive MHT (150 GeV) → $Z(\nu\nu)H$

Expected Sensitivity



Expected sensitivity for exclusion in bb and tt similar and roughly a factor 2 worse than other major modes

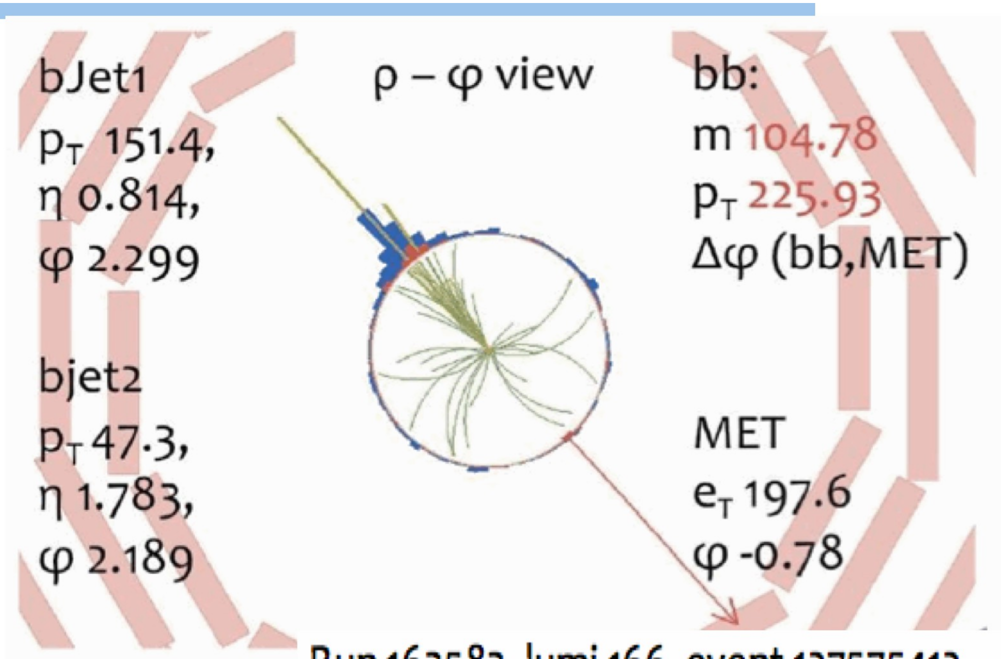
Expected p-values



P-value for all modes

VH Analysis in a nutshell

- 5 modes under study:
  Z(ll)H, W(lν)H, Z(νν)H, l= e,μ



Run 163583 lumi 166 event 127575412
ZvvHbb candidate

- Boosted analysis:
  → Require high momentum vector boson and 2-b tagged jets, back-to -back
  → Better signal to background ratio
  → Two p_T(V) bins

- Use Data control regions to constrain most Important backgrounds (V+jet, Light or Heavy, ttbar)

- b-jet energy regression
  → Mass resolution improvement

- Boosted Decision Tree algorithm (BDT) to discriminate signal versus background

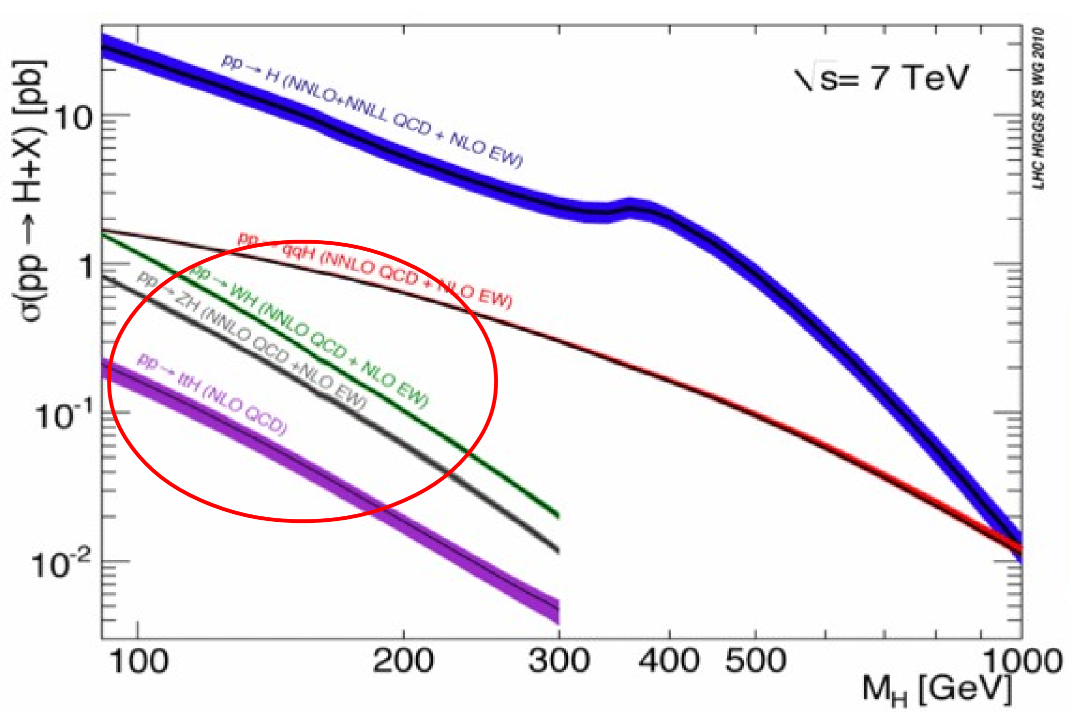
Channel	Medium boost	High boost
ZllH	50<Zpt<100	Zpt>100
WlnH	120< Wpt<170	Wpt>170
ZnnH	120<Zpt<160	Zpt>160

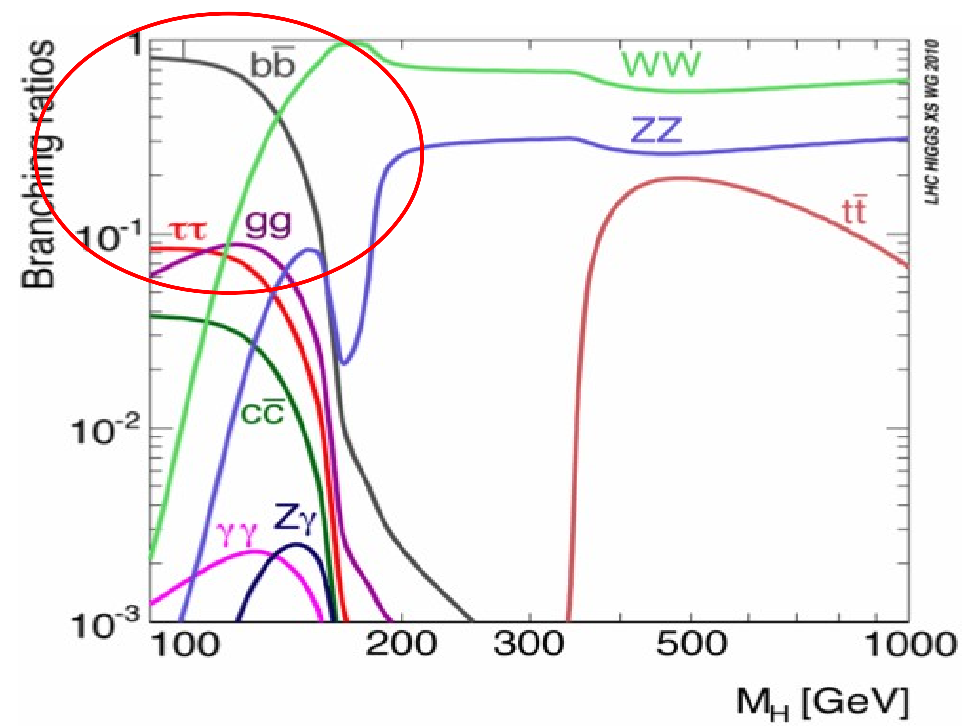
$H \rightarrow b\bar{b}$ and the Higgs Hunting

Given the observation of a new particle at 125 GeV, confirm or Exclude it's the Standard Model Higgs

→ need complementary information from as many channels as possible
→ H → bb largest Branching Ratio by far below 130 GeV

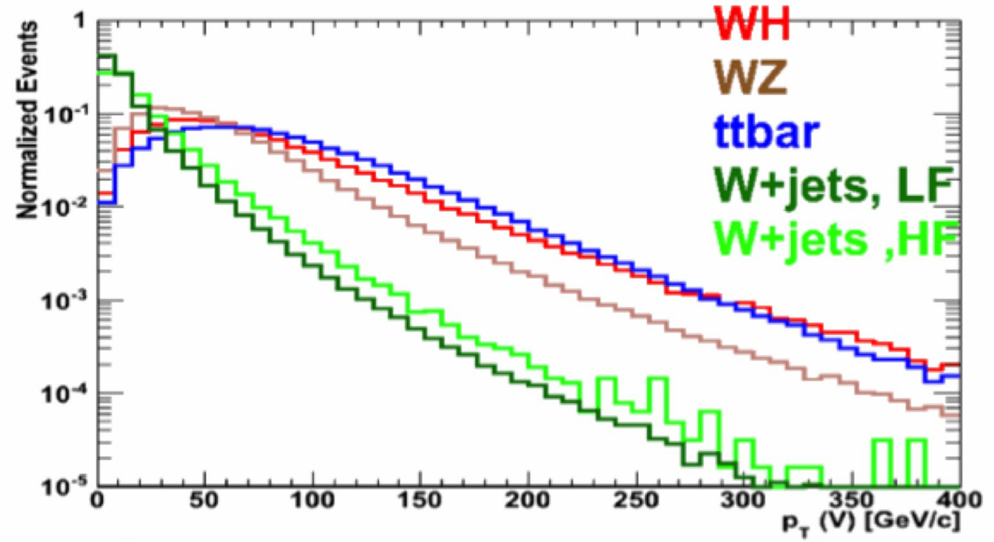
→ Crucial piece in the observation puzzle

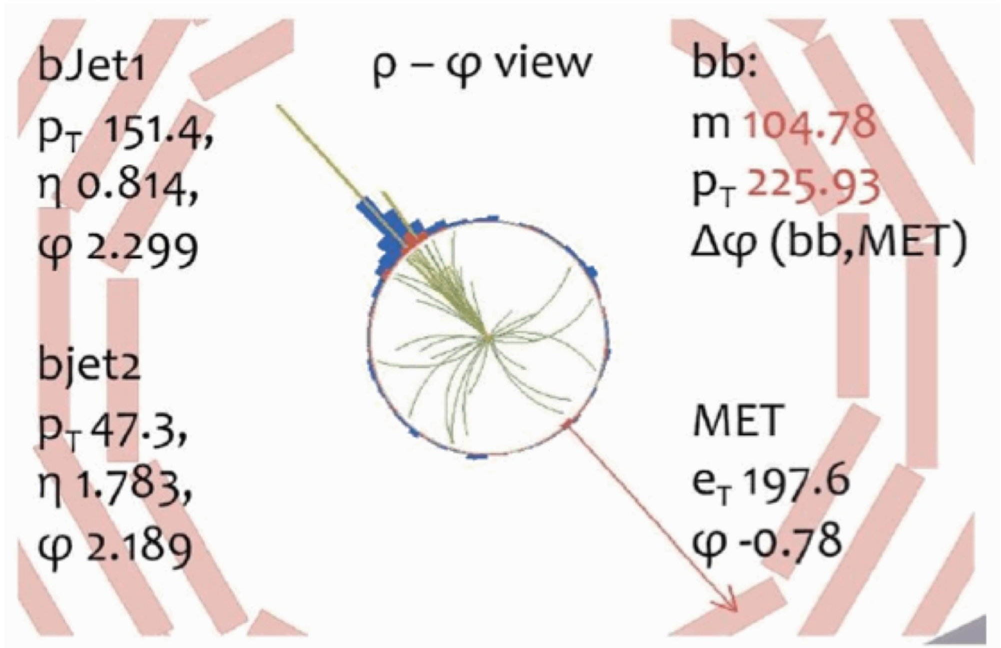




Analysis Strategy

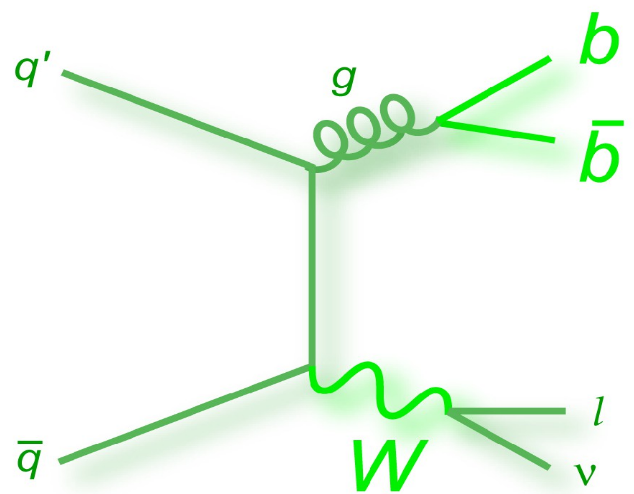
- Enormous background in $H \rightarrow bb$ due to QCD:
  $pp \rightarrow H \rightarrow bb$ deemed impossible
- Use $pp \rightarrow VH$ (V=W,Z) with leptonic V decays
  require high momentum: 'boosted' analysis
- General strategy:
  → boosted vector boson,
  → 2 b-tagged jets,
  → back-to-back

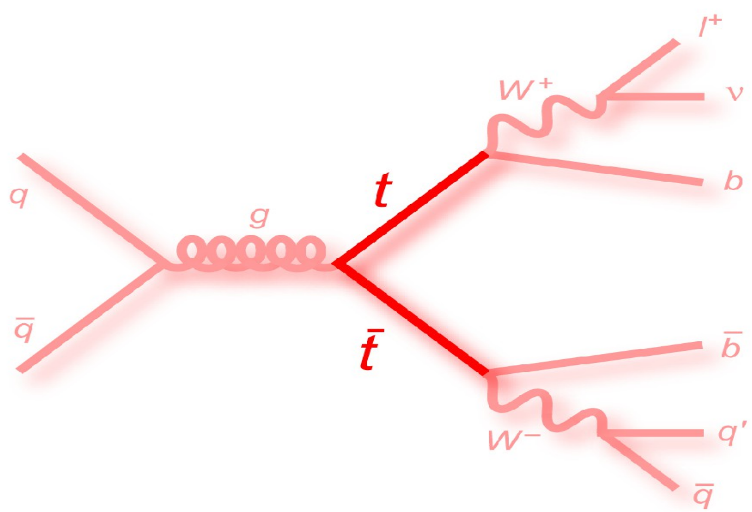




Run 163583 lumi 166 event 127575412
ZννHbb candidate

Backgrounds





Reducible backgrounds
QCD (strongly suppressed by iso and boost)
V+udscg,V+$b\bar{b}$ @ low p_T and mass
W(lν)W(jj)
ttbar and single top (Wb)

Irreducible backgrounds
V+$b\bar{b}$ @ high p_T and mass
ZZ($b\bar{b}$), W(lν)Z($b\bar{b}$)

Important discriminating variables
Mass resolution (separation of VH from VV)
b-tagging (suppression of V+light)
Back-to-back topology
Additional jet activity

Physics Objects (2011)

- Particle Flow based Analysis
  PileUp removal using PFNoPU
  PV selected as the one with highest activity

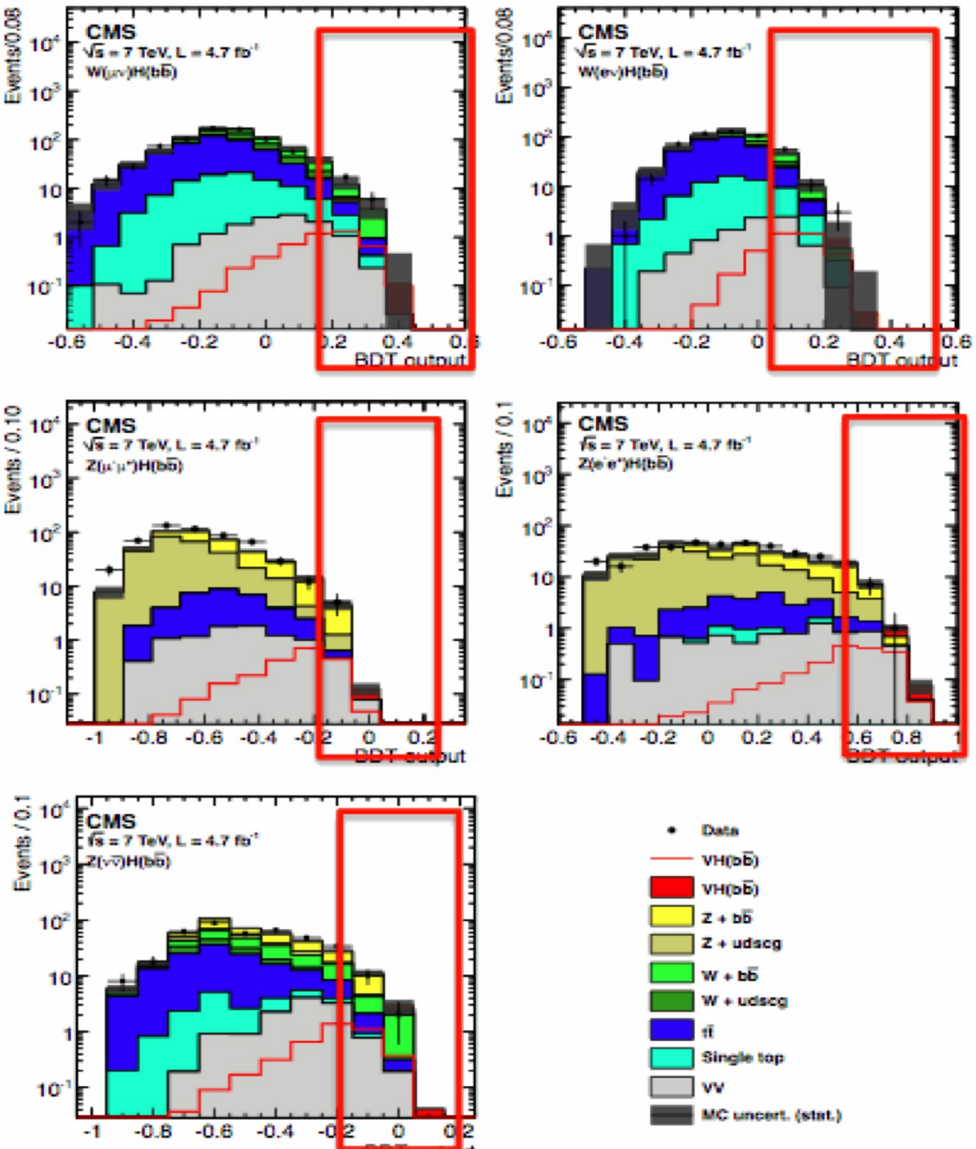
	$Z \to \ell\ell$	$W \to \ell\nu$	$Z \to \nu\nu$	$Z \to \ell\ell$	$W \to \ell\nu$	$Z \to \nu\nu$
Physics Object	p_T (GeV)			ID,Iso		
PF Muon	20, $\lvert\eta\rvert$ <2.4	20, $\lvert\eta\rvert$ <2.4	-	VBTF, PFiso < 0.15		-
PF Electron	20, $\lvert\eta\rvert$ < 2.5, NoGap	30, $\lvert\eta\rvert$ < 2.5, NoGap	-	WP95	WP80	-
AK5 PF Jets	20, $\lvert\eta\rvert$ <2.4	30, $\lvert\eta\rvert$ <2.4	80/30, $\lvert\eta\rvert$ <2.4	Loose		Tight
PFMET	-	35 ($W \to e\nu$)	160	-	-	-
$p_T(V, H)$	100	150-165	160	-	-	-

- MC re-weighted to match PU distribution on data
- Z($\ell\ell$): 75 < m($\ell\ell$) <105 GeV,
- Z($\nu\nu$): PFMET cut and lepton veto
- W($\ell\nu$): Combine PFMET and lepton
  No additional leptons

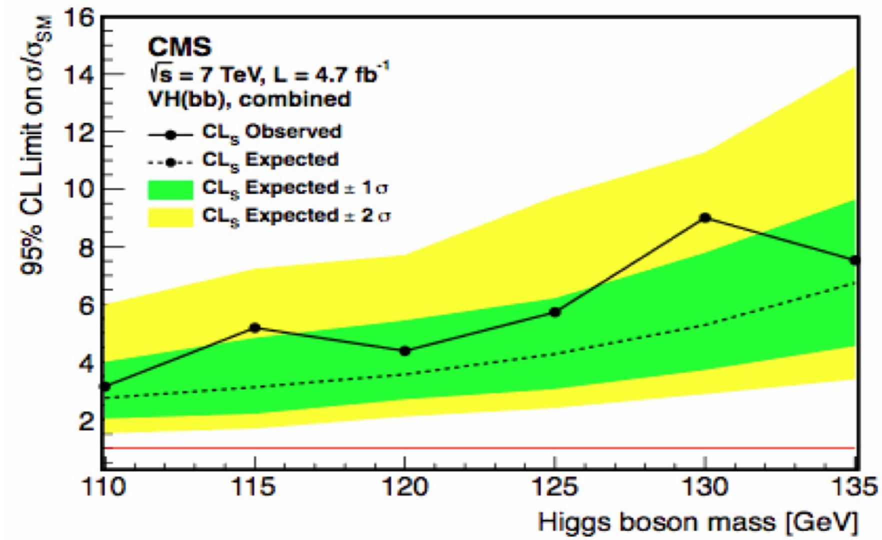
Muon selection:

- Global and Tracker;
- $\chi^2/\mathtt{ndof} < 10$ for the global muon fit;
- Tracks associated to muons must satisfy:
  - at least one pixel hit
  - at least ten total hits (strip + pixel)
  - at least one valid hit in the muon chambers
  - at least two muon stations
  - impact parameter in the transverse plane $d_{xy} < 2\,\mathrm{mm}$

VHbb 2011 Results



m_H (GeV)	110	115	120	125	130	135
BDT Exp.	2.7	3.1	3.6	4.3	5.3	6.7
BDT Obs.	3.1	5.2	4.4	5.7	9.0	7.5
$m(jj)$ Exp.	3.0	3.2	4.4	4.7	6.4	7.7
$m(jj)$ Obs.	3.4	5.6	6.7	6.3	10.5	8.9



Final yield estimate based on Cut and Count on the BDT discriminant

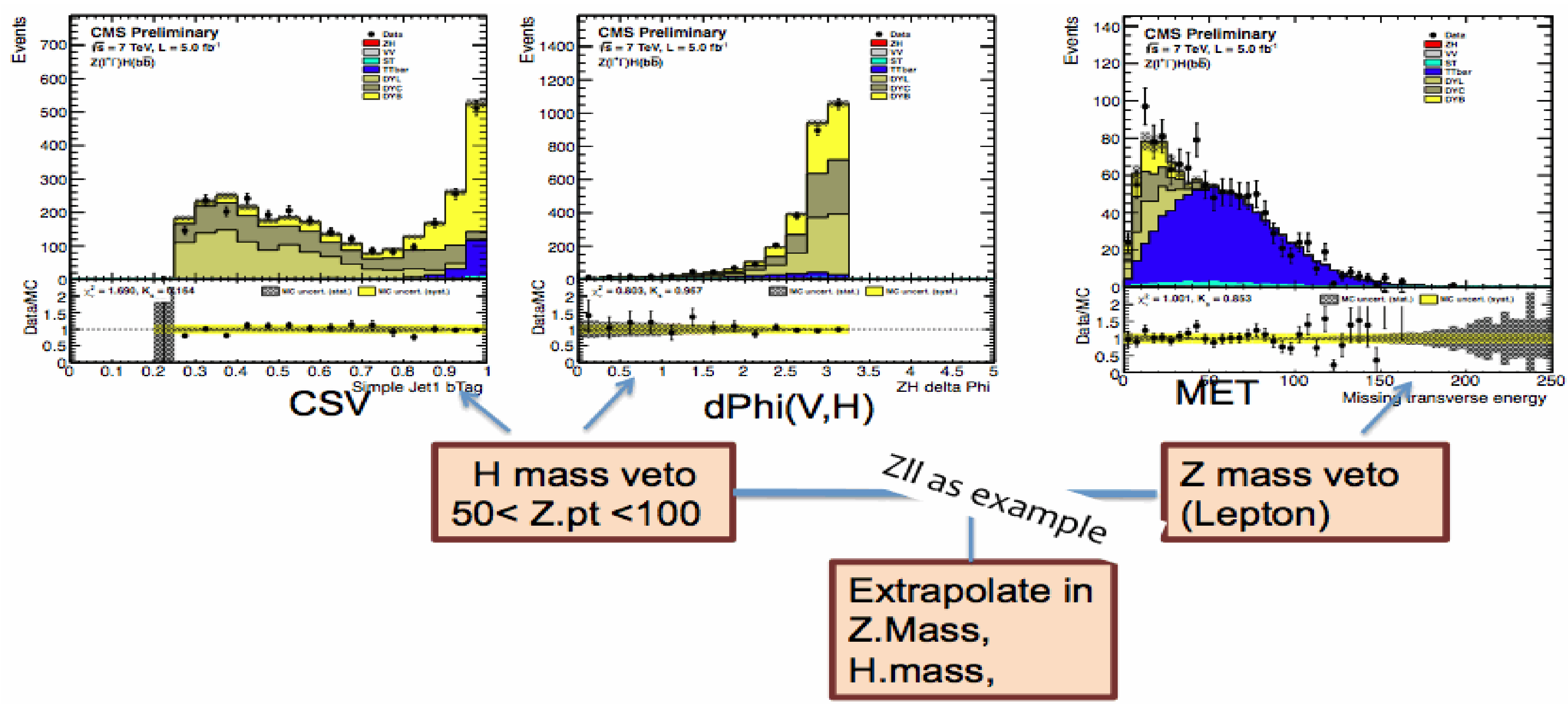
Simple Cut and Count analysis on di-jet invariant Mass (MJJ) as a cross-check

PLB 710(2012)284-306

2011 Improvements

Category	2011	ICHEP 2012	Sensitivity Gain
Background Treatment	Event Count in Control Regions	Fit shapes in Control Regions	
Higgs Reconstruction	AK5PF di-jet with standard corrections	Regression	10-20%
Boost	Single bin, high boost analysis	Two bins (add medium boost)	10%
BDT && MJJ	Cut and Count	Shape Analysis	20%

Control Region Shape Fit



- Scale Factors for V+light(heavy) and ttbar background re-weighting extracted from simultaneous binned Maximum Likelihood fit in 3 control regions
- Control regions defined as kinematically close to Signal Region, still independent

Background Scale Factors

- Scale factors for background re-weighting largely consistent between 7 and 8 TeV analysis

Process	WH	$Z(\ell\ell)H$	$Z(\nu\nu)H$
Low p_T			
W + udscg	$0.88 \pm 0.01 \pm 0.03$	–	$0.89 \pm 0.01 \pm 0.03$
$Wb\bar{b}$	$1.91 \pm 0.14 \pm 0.31$	–	$1.36 \pm 0.10 \pm 0.15$
Z + udscg	–	$1.11 \pm 0.03 \pm 0.11$	$0.87 \pm 0.01 \pm 0.03$
$Zb\bar{b}$	–	$0.98 \pm 0.05 \pm 0.12$	$0.96 \pm 0.02 \pm 0.03$
$t\bar{t}$	$0.93 \pm 0.02 \pm 0.05$	$1.03 \pm 0.04 \pm 0.11$	$0.97 \pm 0.02 \pm 0.04$
High p_T			
W + udscg	$0.79 \pm 0.01 \pm 0.02$	–	$0.78 \pm 0.02 \pm 0.03$
$Wb\bar{b}$	$1.49 \pm 0.14 \pm 0.19$	–	$1.48 \pm 0.15 \pm 0.20$
Z + udscg	–	$1.11 \pm 0.03 \pm 0.11$	$0.97 \pm 0.02 \pm 0.04$
$Zb\bar{b}$	–	$0.98 \pm 0.05 \pm 0.12$	$1.08 \pm 0.09 \pm 0.06$
$t\bar{t}$	$0.84 \pm 0.02 \pm 0.03$	$1.03 \pm 0.04 \pm 0.11$	$0.97 \pm 0.02 \pm 0.04$

7 TeV Analysis

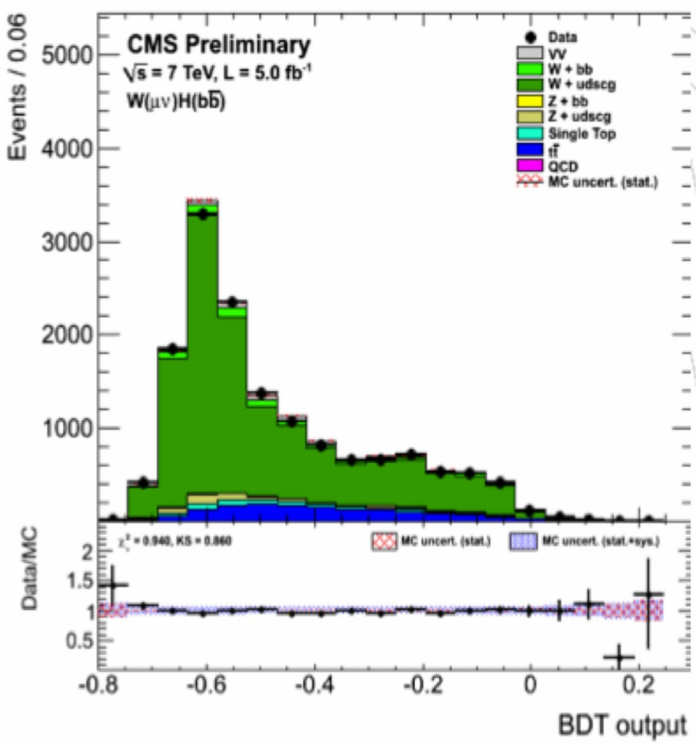
Process	WH	$Z(\ell\ell)H$	$Z(\nu\nu)H$
Low p_T			
W + udscg	$0.97 \pm 0.01 \pm 0.03$	–	$0.96 \pm 0.04 \pm 0.03$
$Wb\bar{b}$	$2.0 \pm 0.24 \pm 0.32$	–	$1.48 \pm 0.34 \pm 0.151$
Z + udscg	–	$1.33 \pm 0.03 \pm 0.10$	$0.96 \pm 0.05 \pm 0.03$
$Zb\bar{b}$	–	$1.14 \pm 0.05 \pm 0.14$	$0.92 \pm 0.10 \pm 0.050$
$t\bar{t}$	$1.12 \pm 0.02 \pm 0.05$	$1.02 \pm 0.04 \pm 0.11$	$1.02 \pm 0.035 \pm 0.03$
High p_T			
W + udscg	$0.87 \pm 0.01 \pm 0.03$	–	$0.85 \pm 0.04 \pm 0.03$
$Wb\bar{b}$	$1.30 \pm 0.23 \pm 0.13$	–	$1.48 \pm 0.25 \pm 0.20$
Z + udscg	–	$1.33 \pm 0.03 \pm 0.10$	$1.052 \pm 0.04 \pm 0.04$
$Zb\bar{b}$	–	$1.14 \pm 0.05 \pm 0.14$	$1.13 \pm 0.07 \pm 0.08$
$t\bar{t}$	$0.97 \pm 0.02 \pm 0.04$	$1.02 \pm 0.04 \pm 0.11$	$1.01 \pm 0.05 \pm 0.04$

8 TeV Analysis

- Uncertainties include: MC statistics, detector effect (jet resolution and scale, b-tag efficiency and mis-id) and estimated by repeating the fit with template variations

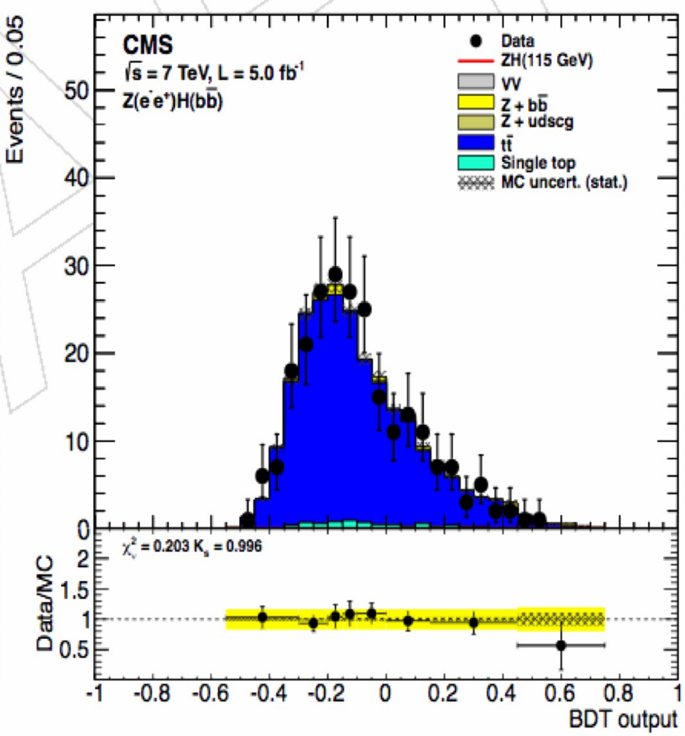
BDT Test In Control Regions

W(μν)H



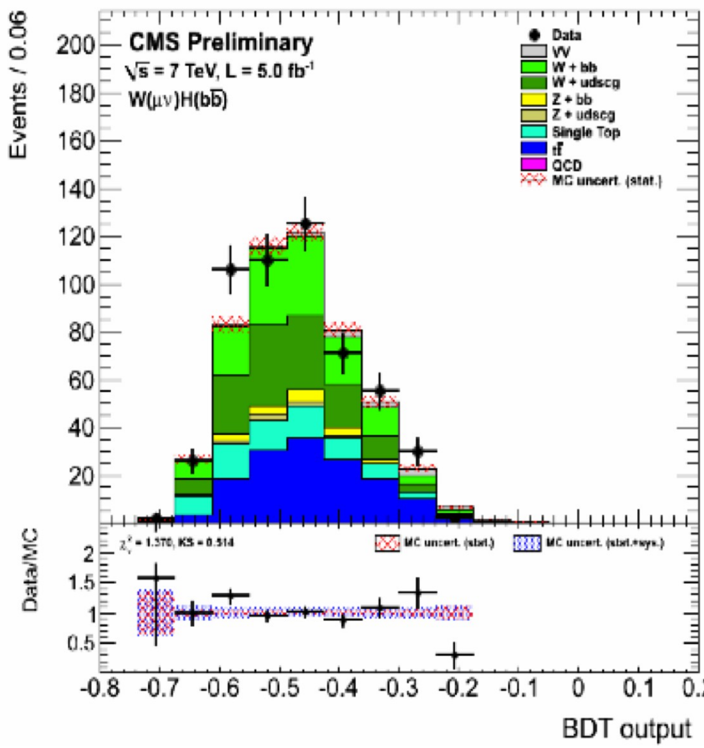
W+light

Z(ee)H



ttbar

W(μν)H



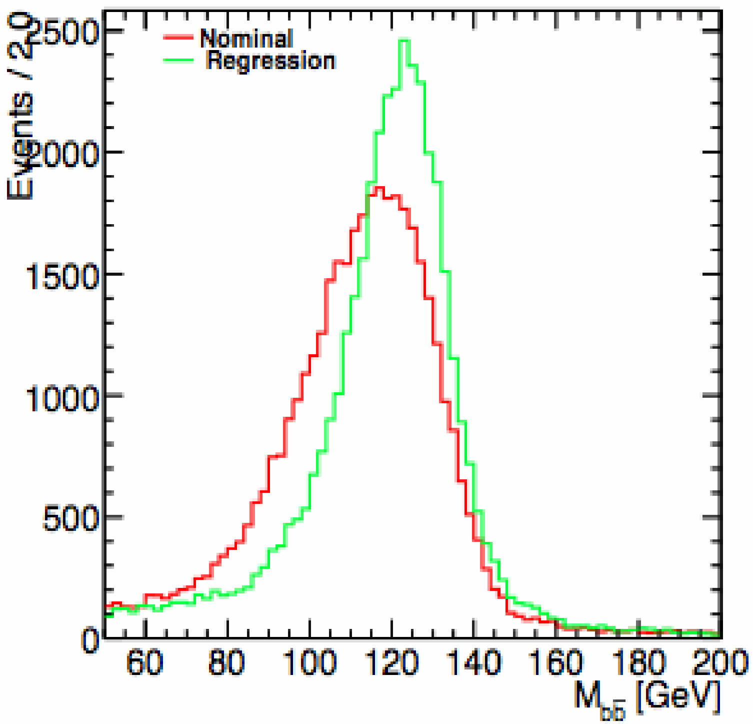
W+heavy

Excellent agreement of BDT output in different kinematic regions and background composition proves BDT robustness

B-jet energy Regression

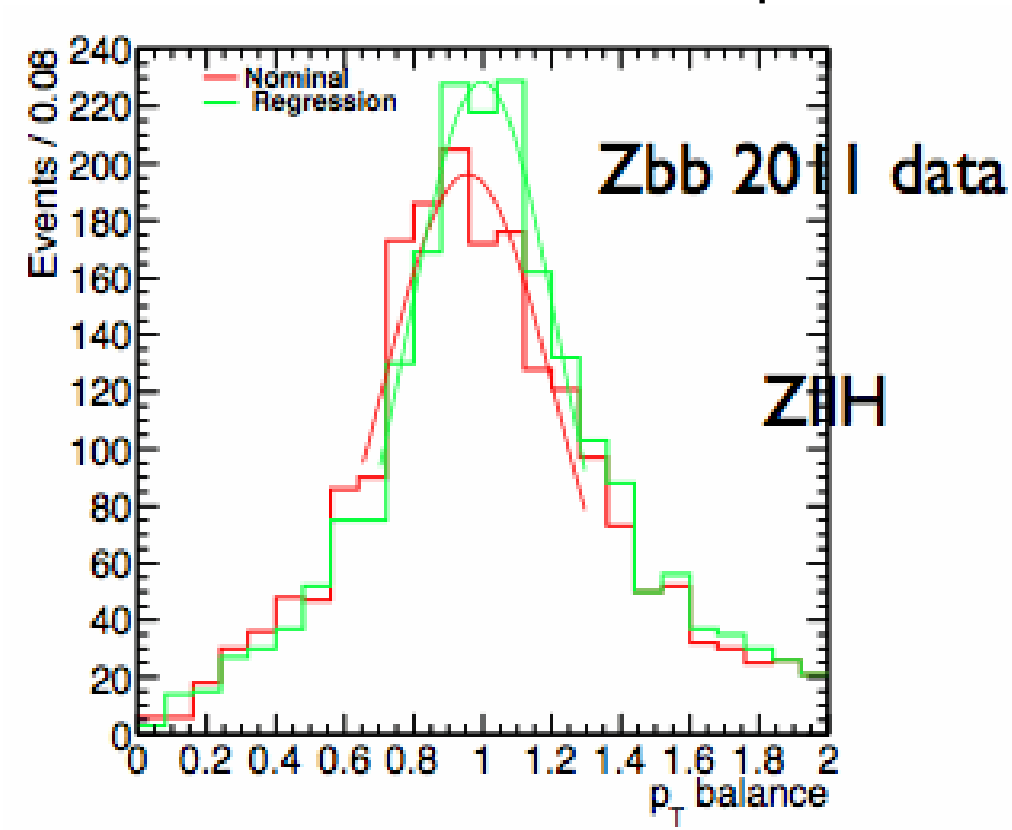
New since 2011 Analysis

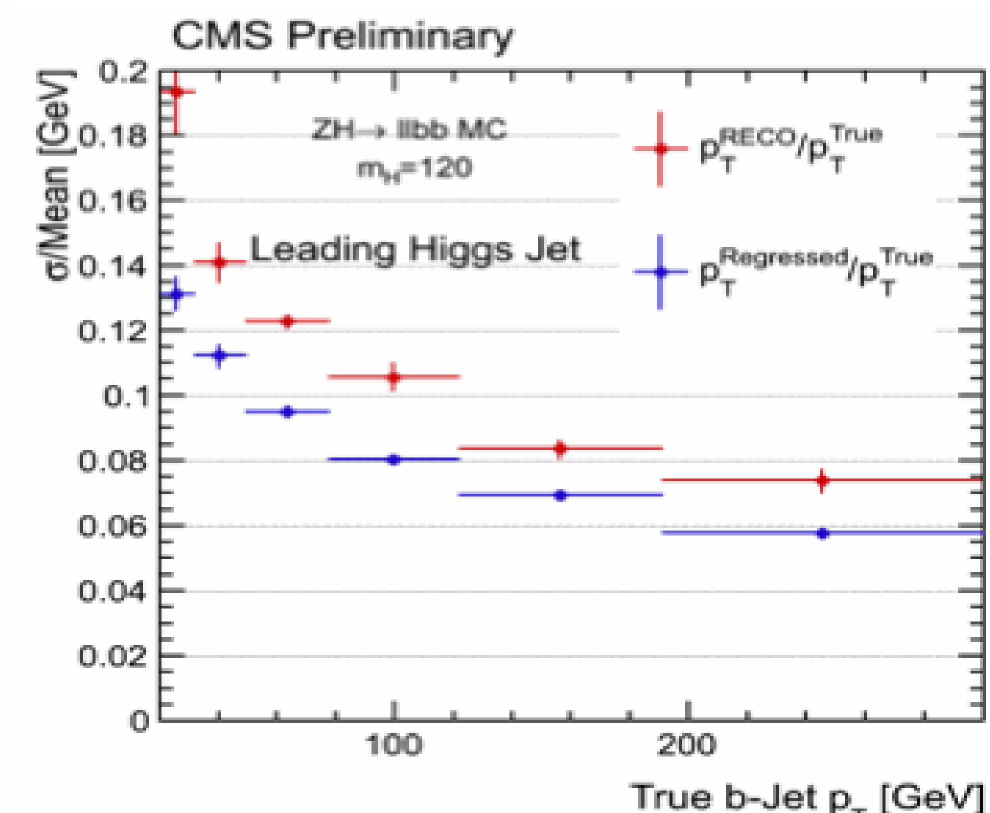
- Implementation based on NN method developed at CDF for b-jet energy corrections: http://arxiv.org/pdf/1107.3026.pdf
- Multivariate Regression (BDT) trained on VH signal events using several (b)-jet variables
  → p_T, η, Uncorrected p_T, E_T, M_T, p_TLeadTrack, charged had fraction, Secondary Vertex info (if any) MET in Z(ll)H events
  → Training at all mass points simultaneously to avoid mass bias
- Improvements in resolution of the order of 20% for Z(ll)H, 15%for W(lν)H and Z(νν)



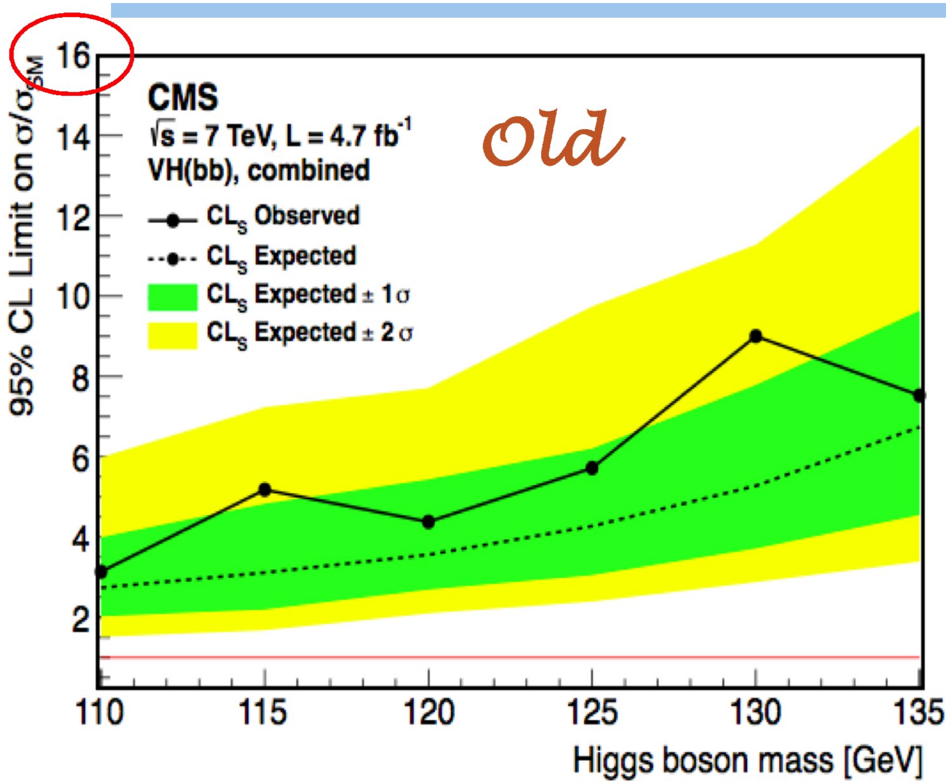
Regression Validation

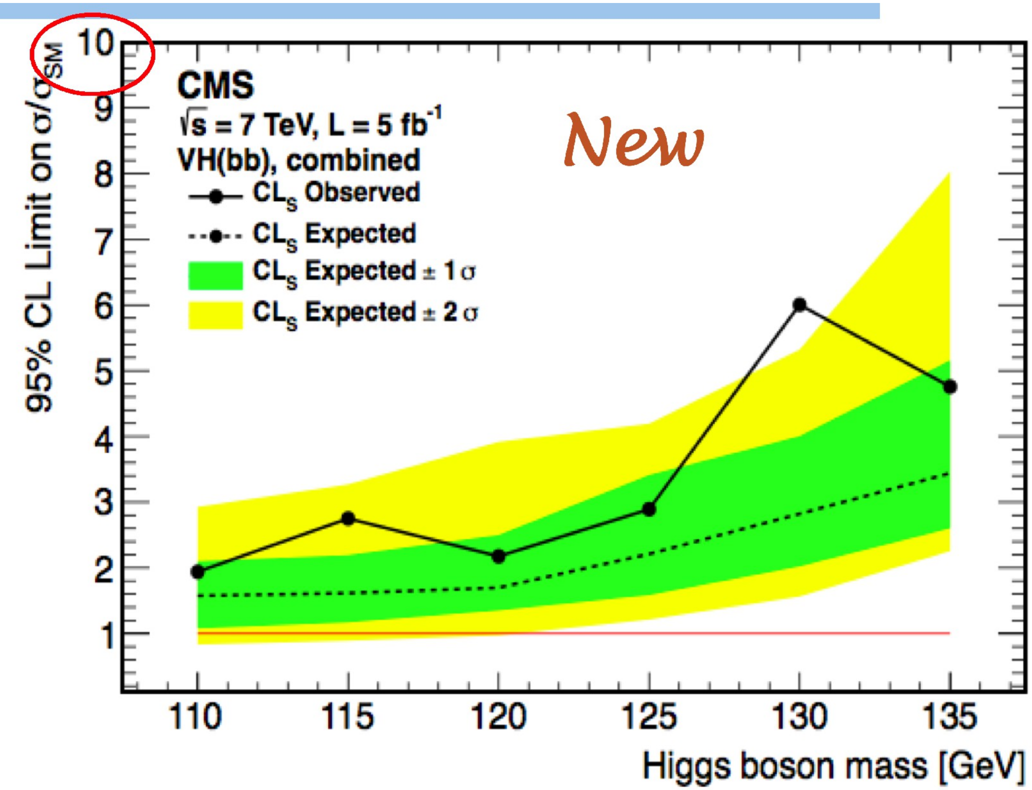
- Extensively validated on simulation and Data Control Regions
  → check of data/MC agreement of variables input to the regression in all control regions
  → p_T balance in Z(ll)+bb
  → full reconstruction of top mass in ttbar and Single Top samples





SM Exclusion Limits (2011)





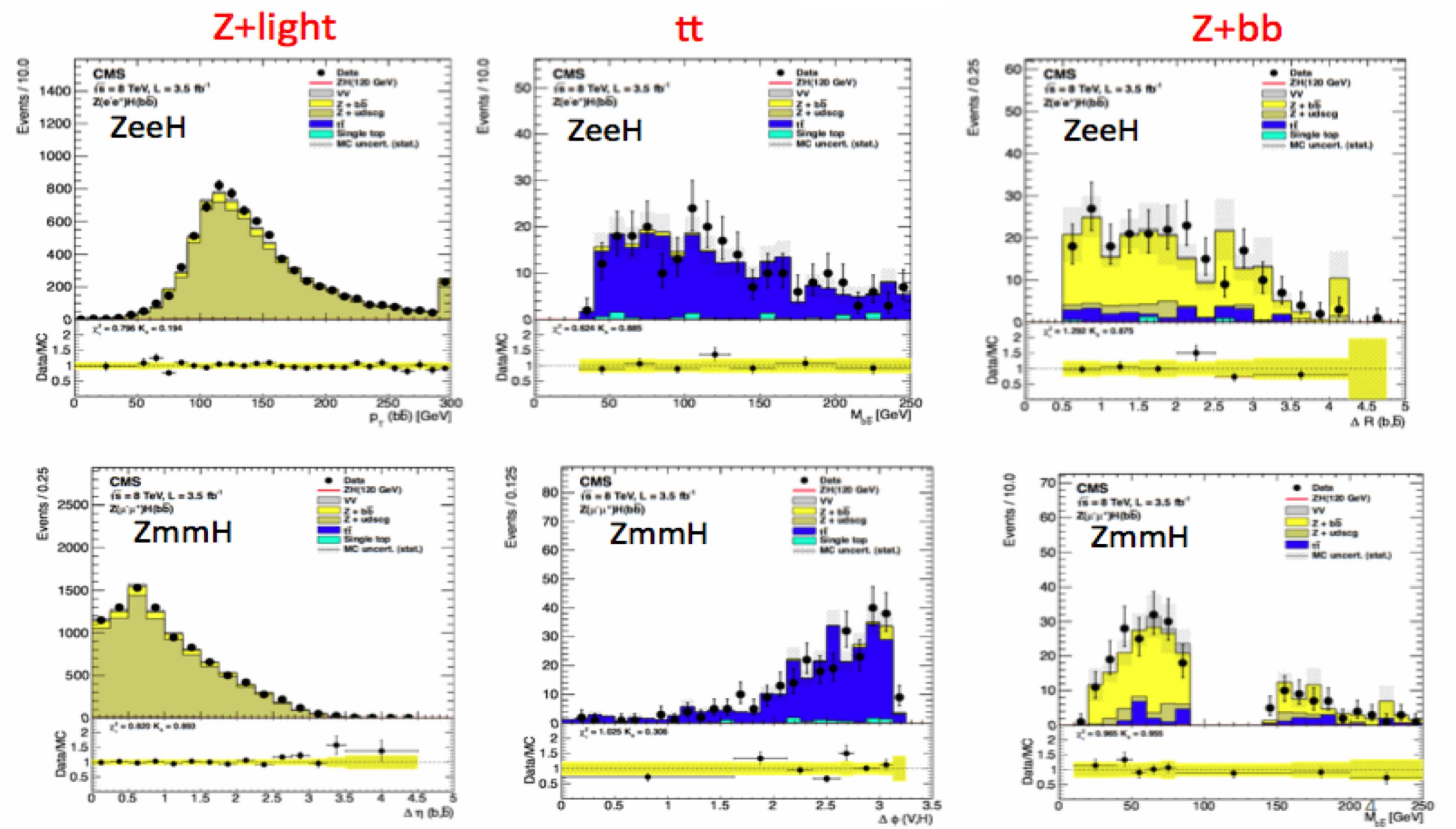
Expected limit improves by ~50%

Broad excess, 115-135 GeV

	110	115	120	125	130	135
Exp	1.57	1.61	1.69	2.21	2.82	3.44
Obs	1.93	2.75	2.17	2.89	6.0	4.8

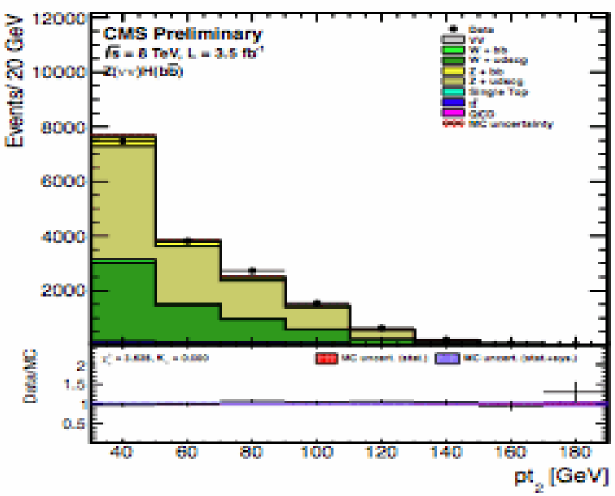
Shape of the observed limit very similar compared to published analysis

Control Regions Data/MC

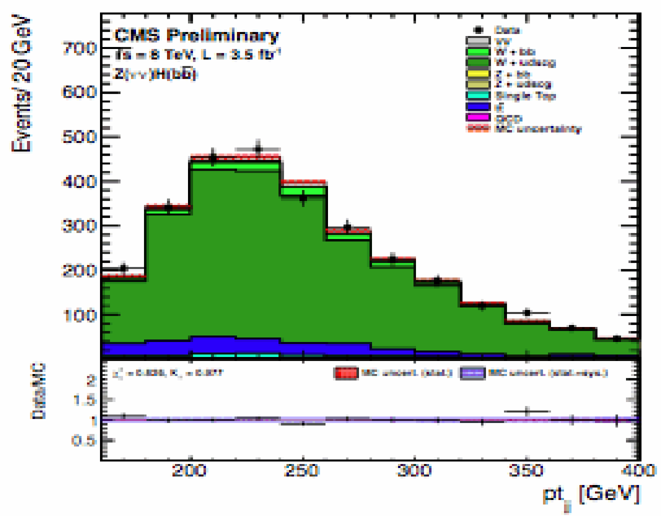


Control Regions Data/MC

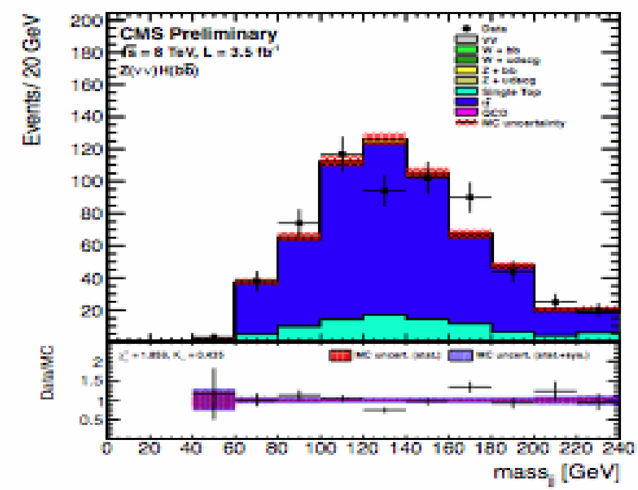
Z+light



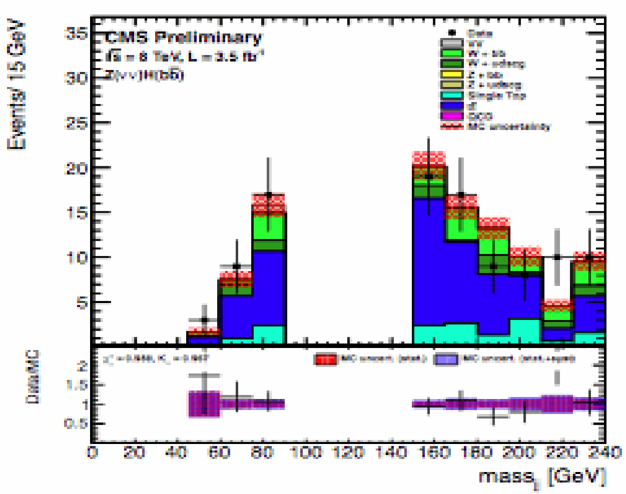
W+light

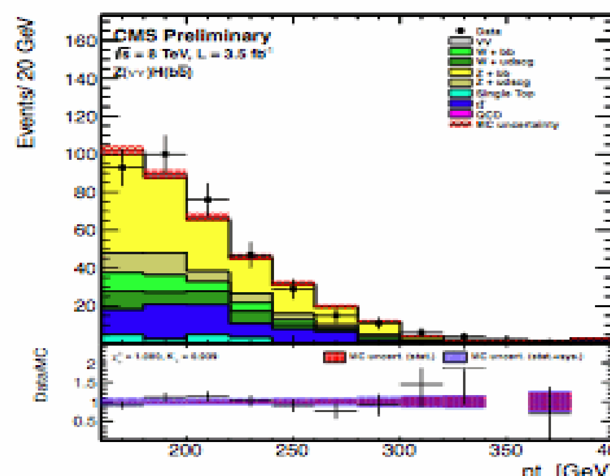


tt



W+bb





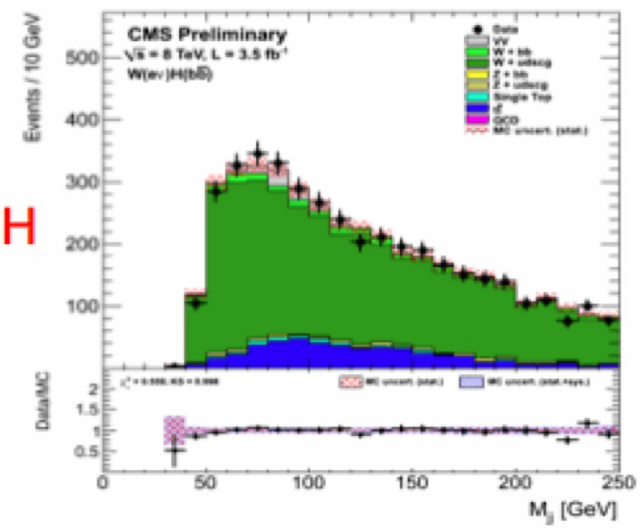
Z+bb

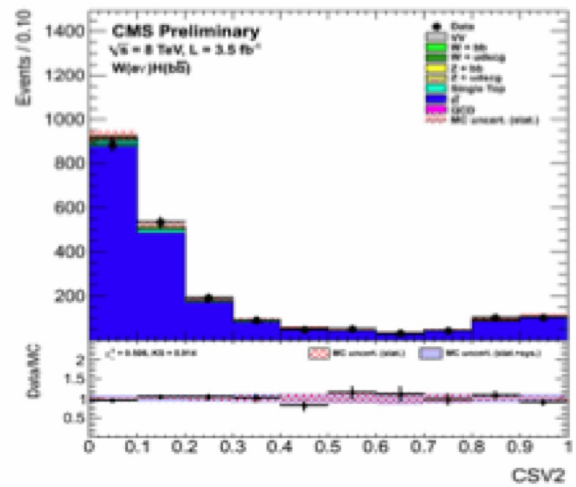
Good agreement across the board

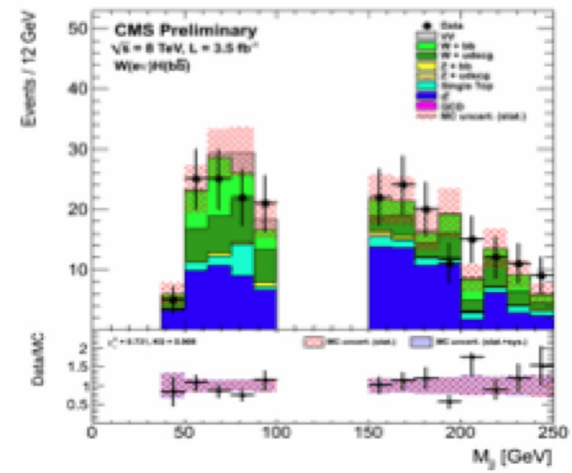
Control Regions Data/MC

W+light | $t\bar{t}$ | W+bb

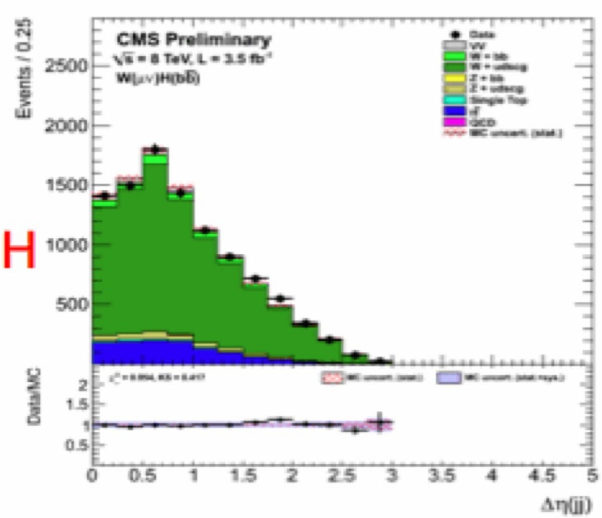
WeH

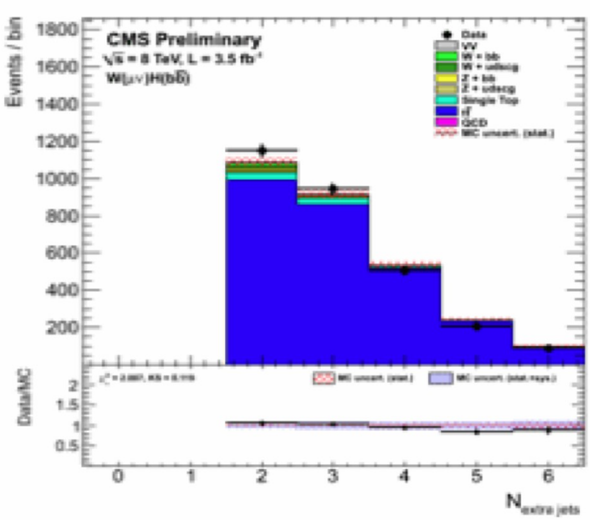


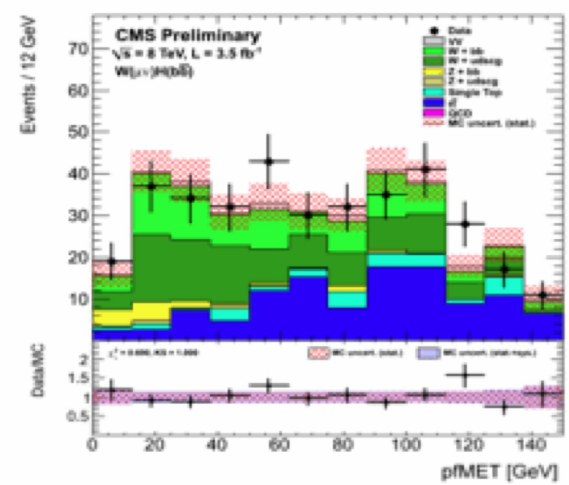




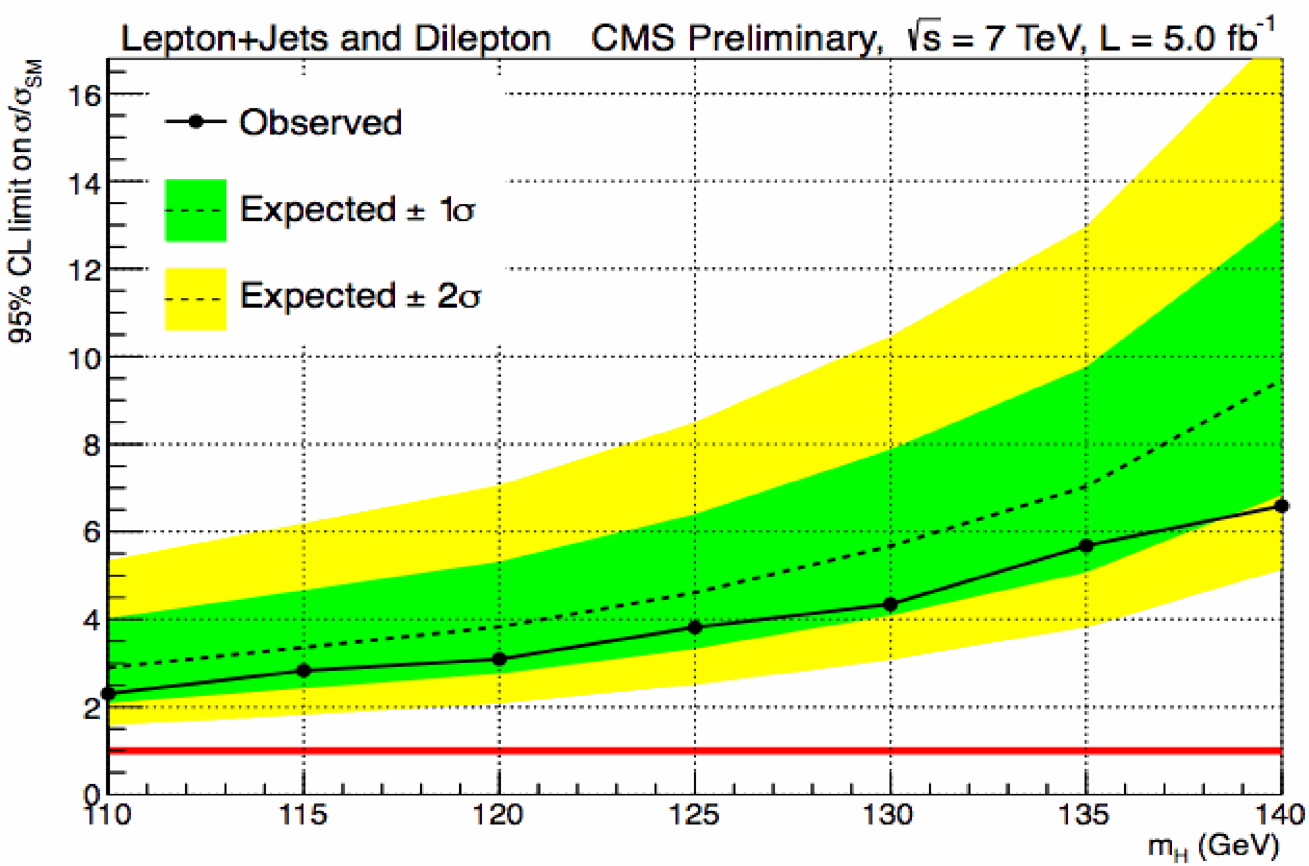
WmH







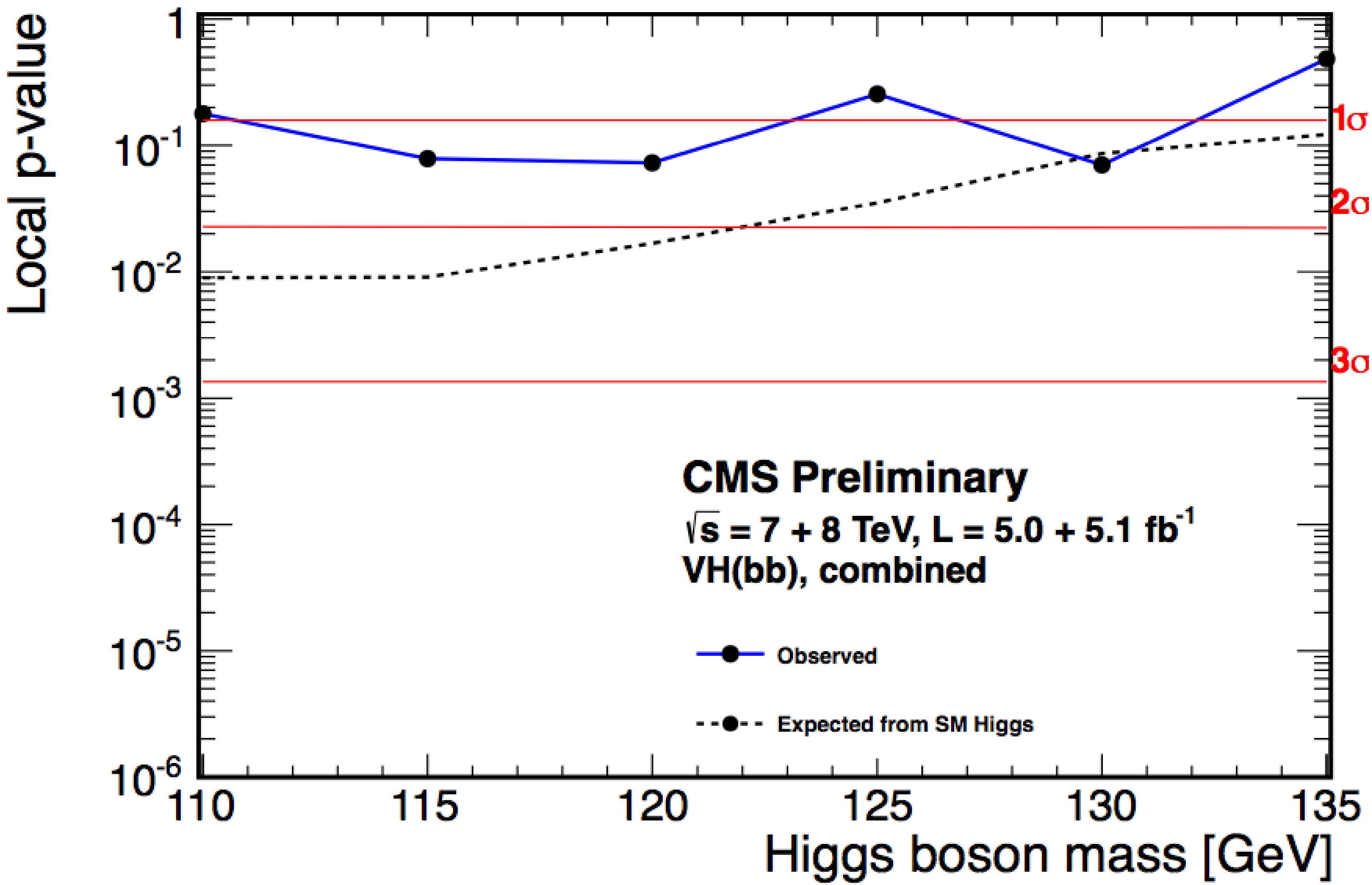
ttH Exclusion Limits



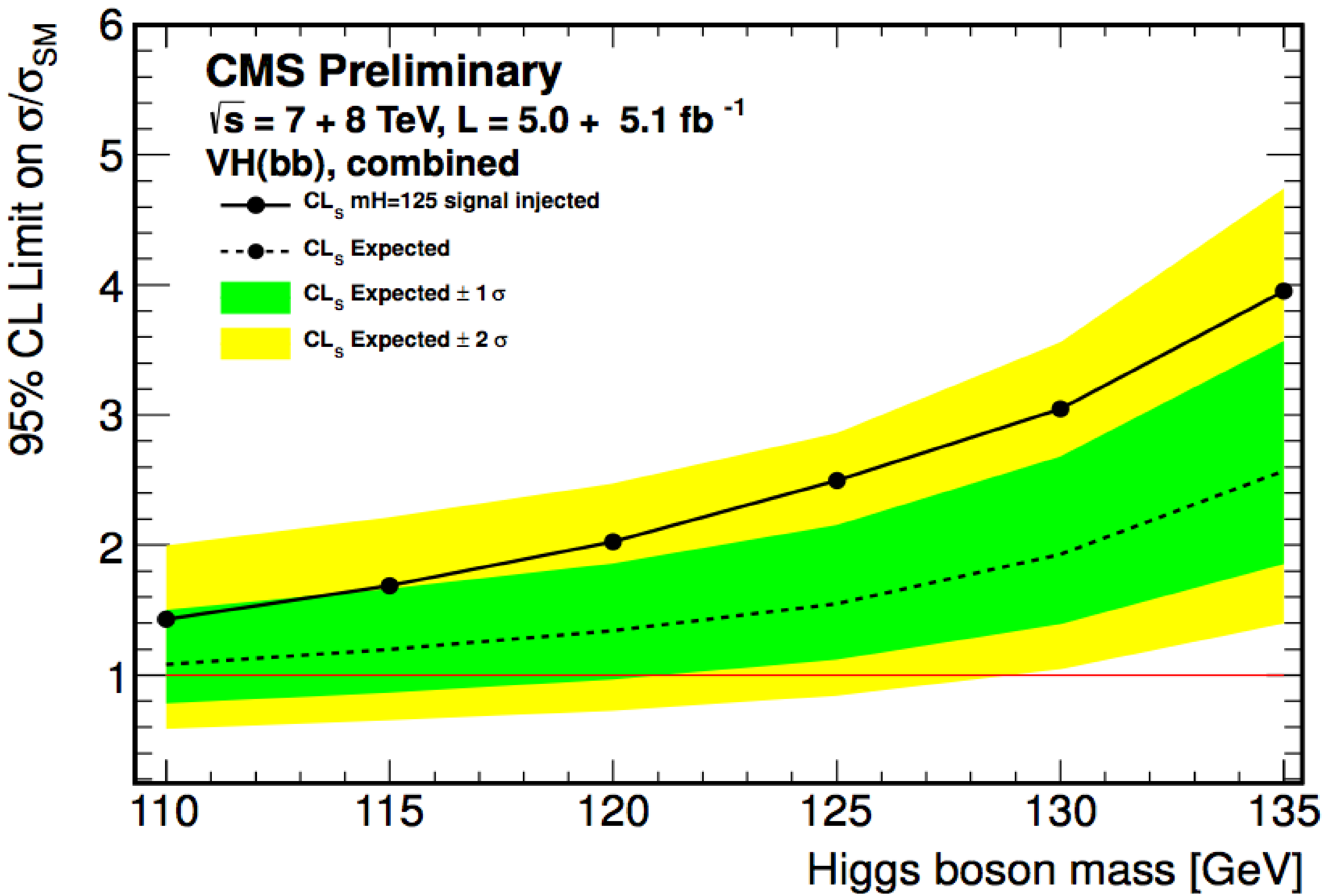
Mass	Exp.	Obs.
110	2.90	2.30
115	3.36	2.83
120	3.83	3.09
125	4.61	3.82
130	5.67	4.35
135	7.03	5.68
140	9.47	6.59

- Sensitivity dominated by lepton+jet mode, 5-10% improvement from di-lepton mode
- Dominant uncertainties: b-tag, JES in LJ, factorization scale in DIL
- No excess seen, expect 4.6 x σ_{SM} at 125 GeV, observe 3.8 x σ_{SM}

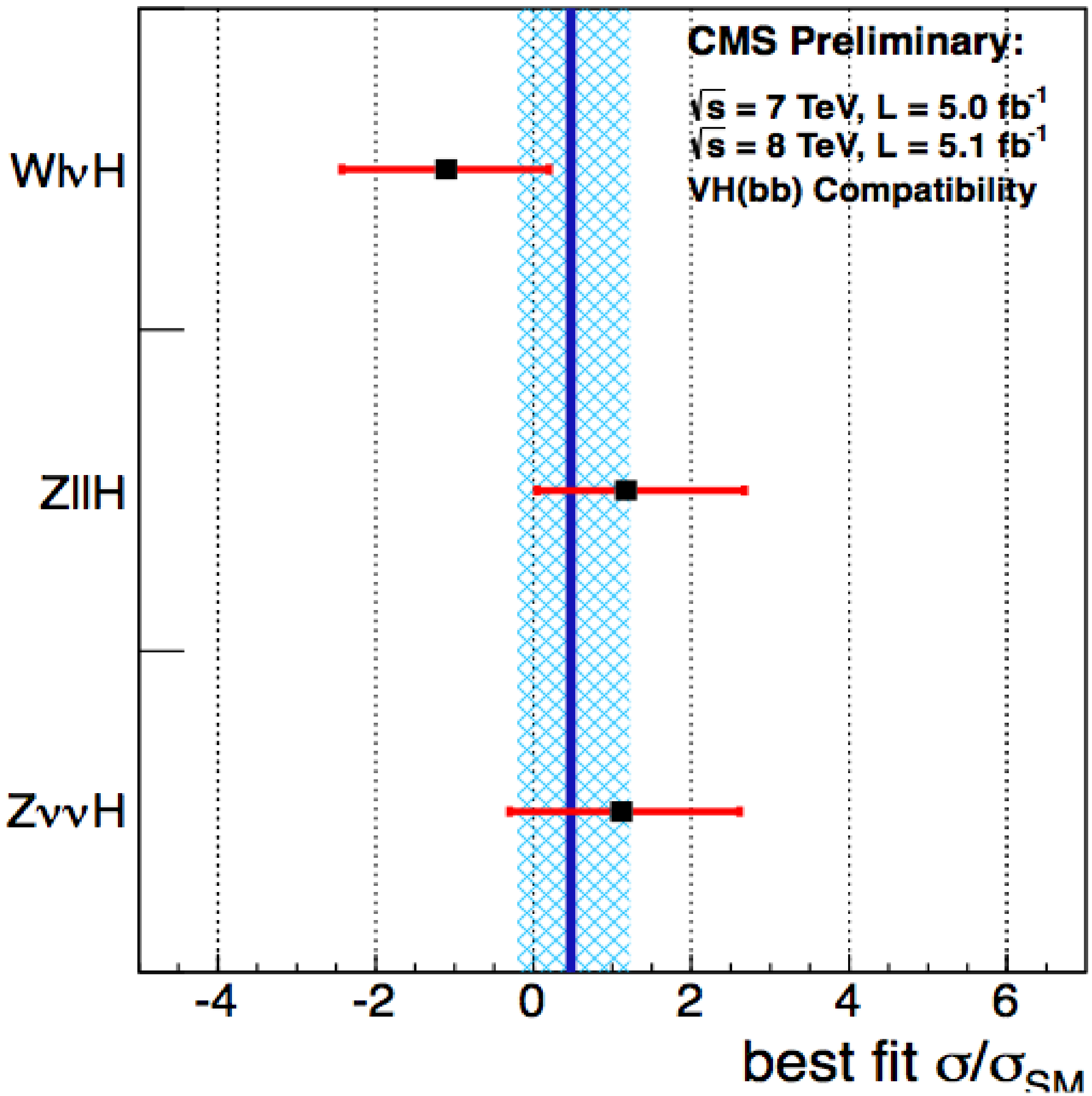
P-values (7+8 TeV)



MH(125) signal injection

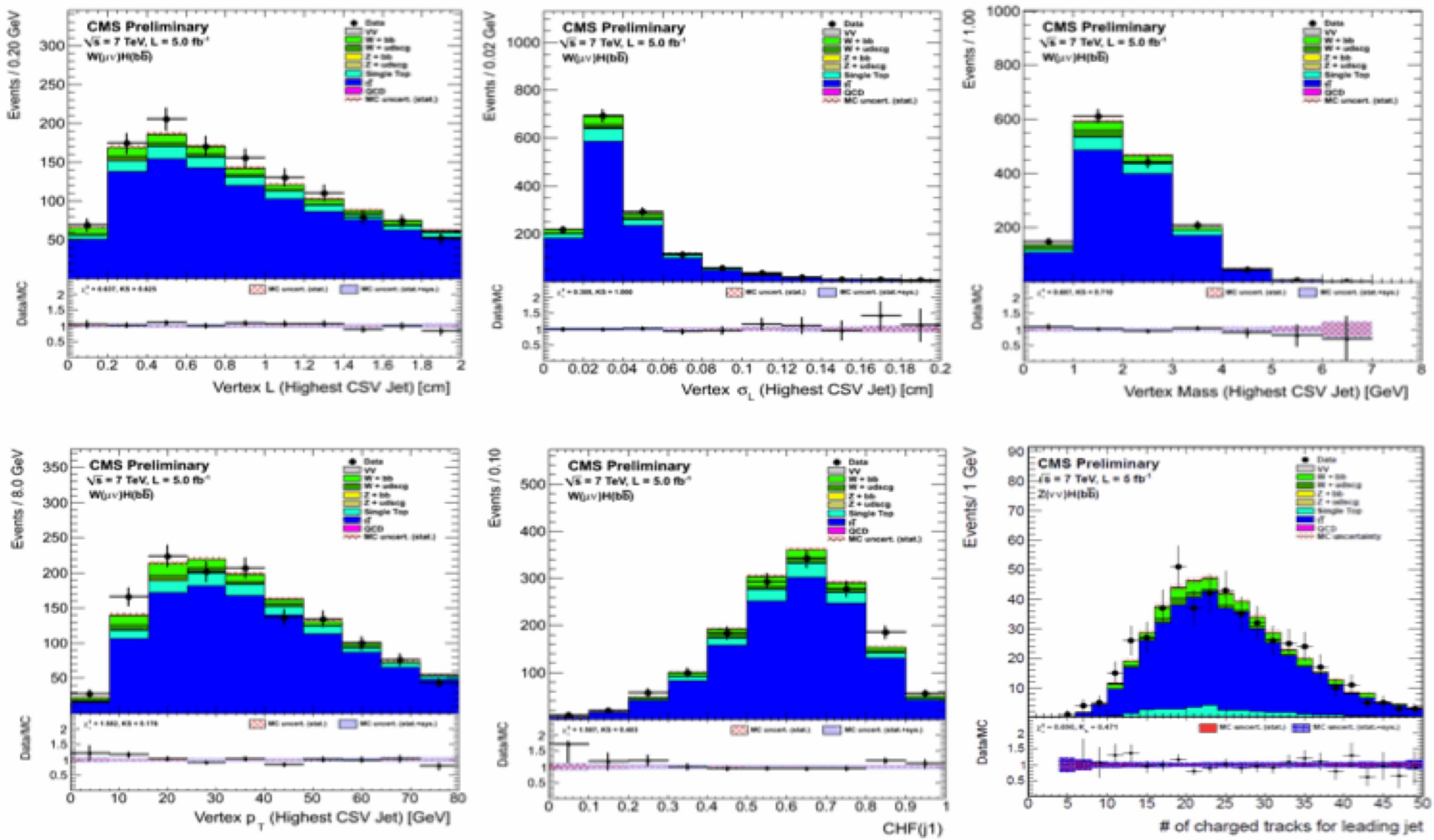


σ/σ_SM compatibility

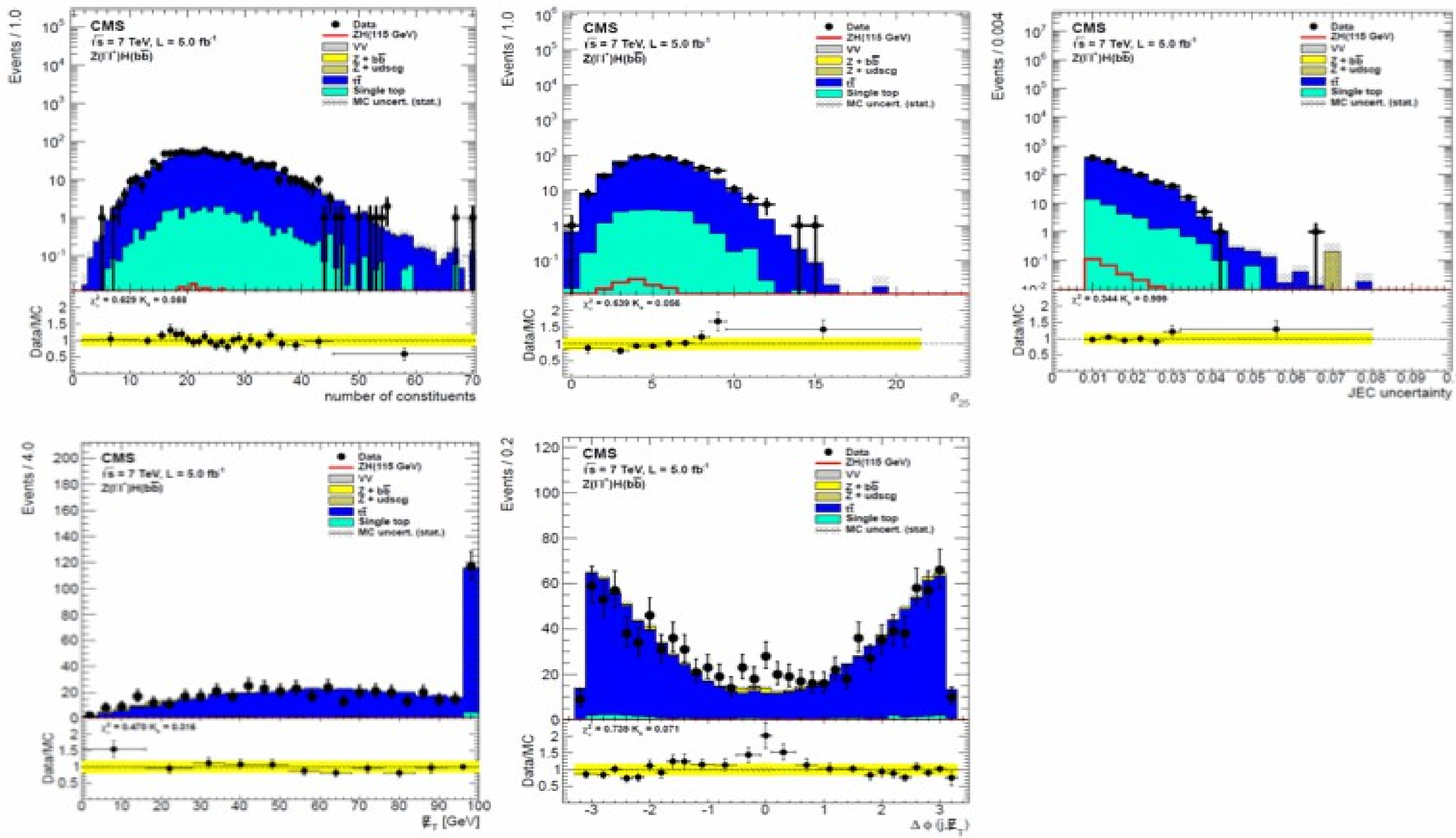


7 TeV
+ 8 TeV

Regression Input Variables



Regression Input Variables



Systematic Uncertainties

- Signal:
  Higgs cross-section: use NNLO from LHC WG, currently estimate 4% error (PDF+alphas, scale)

  p_T spectrum: recent theoretical calculations address our boosted regime: 5(10)% for Z(W)H due to electroweak corrections (http://arxiv.org/abs/0710.4749) and 10% from QCD (NNLO vs NLO, http://www.arxiv.org/abs/1107.1164)

- Background:
  Data-driven:
  Uncertainty on the SF determination →
  1) Statistical uncertainty
  2) systematic on CR definition
  From CR: V+jets (light: 7%, heavy: 16%), ttbar (8%)

  MC based: VV (30%), single top (30%)

MJJ/BDT Cut Efficiency

Variable	$W(\mu\nu)H$	$W(e\nu)H$	$Z(\mu\mu)H$	$Z(ee)H$	$Z(\nu\nu)H$
Pre-select	10.68 ± 0.08	5.845 ± 0.053	11.98 ± 0.61	10.73 ± 0.04	15.13 ± 0.08
$p_T(jj)$	14.11 ± 0.27	18.96 ± 0.429	36.35 ± 1.00	37.28 ± 0.21	40.01 ± 0.34
$p_T(V)$	74.83 ± 1.64	76.75 ± 1.990	80.75 ± 1.20	74.80 ± 0.31	–
CSV1	86.96 ± 2.05	62.37 ± 2.012	84.03 ± 1.22	60.14 ± 0.41	58.24 ± 0.66
CSV2	48.69 ± 1.64	60.14 ± 2.454	36.38 ± 2.02	47.54 ± 0.53	48.51 ± 0.79
$\Delta\phi(V,H)$	85.75 ± 2.90	87.17 ± 3.787	88.46 ± 2.15	87.83 ± 0.51	84.93 ± 1.50
N_{aj}	76.41 ± 3.18	73.14 ± 3.704	98.02 ± 2.18	96.07 ± 0.32	80.96 ± 1.59
N_{al}	76.41 ± 3.18	100 ± 5.06	–	–	100
pfMET	–	92.84 ± 4.93	–	–	83.69 ± 1.69
pfMETsig	–	–	–	–	–
$\Delta\phi(\text{pfMET}, J)$	–	–	–	–	92.79 ± 2.07
$M(jj)$	76.93 ± 3.65	82.22 ± 4.81	70.91 ± 2.58	70.20 ± 0.77	75.92 ± 1.94
Total Eff.	0.24 ± 0.01	0.16 ± 0.01	0.66 ± 0.02	0.51 ± 0.01	0.693 ± 0.017

Variable	$W(\mu\nu)H$	$W(e\nu)H$	$Z(\mu\mu)H$	$Z(ee)H$	$Z(\nu\nu)H$
Pre-select	10.67 ± 0.08	5.845 ± 0.053	11.98 ± 0.61	10.73 ± 0.04	15.13 ± 0.08
$p_T(jj)$	18.01 ± 0.30	24.20 ± 0.48	36.35 ± 1.00	37.28 ± 0.21	40.01 ± 0.34
$p_T(V)$	73.20 ± 1.44	73.58 ± 1.71	80.75 ± 1.20	74.80 ± 0.31	–
CSV1	87.06 ± 1.84	87.86 ± 2.16	91.82 ± 1.17	90.80 ± 0.24	31.84 ± 0.31
CSV2	47.70 ± 1.46	50.51 ± 1.71	52.61 ± 1.61	51.94 ± 0.43	40.88 ± 0.62
N_{al}	100	100	–	–	100
pfMET/trg	–	91.95 ± 3.29	–	–	78.24 ± 1.34
BDT	39.82 ± 1.93	36.29 ± 2.13	27.40 ± 2.88	33.45 ± 0.57	52.01 ± 0.67
Total Eff.	0.23 ± 0.01	0.154 ± 0.01	0.47 ± 0.02	0.47 ± 0.01	0.73 ± 0.02

MJJ/BDT Cut Efficiency

Experimental Uncertainties		Propagation into Limit Calculation		
Uncertainty	Uncert.	*0-Jet*	*Boost*	*VBF*
Electron ID & Trigger (*)	±2%	±2%	±2%	±2%
Muon ID & Trigger (*)	±2%	±2%	±2%	±2%
Tau ID & Trigger (*)	±7%	±7%	±7%	±7%
JES (Norm.) (*)	±2.5 – 5%	∓1%	±5%	±10%
b-Tag Efficiency (*)	±10%	∓1%	∓2%	∓2%
Mis-Tagging (*)	±30%	∓1%	∓1%	∓1%
Norm. $Z \to \tau\tau$	±3%	±3%	±5%	±13%
Norm. $t\bar{t}$ (*)	±10 – 30%	±10%	±12%	±30%
Norm EWK	±30%	±30%	±15 – 30%	±30 – 100%
Norm Fakes	±10 – 30%	±10%	±10%	±30%
Lumi (Signal & EWK)	±2.2(5)%	±2.2(5)%	±2.2(5)%	±2.2(5)%
Norm. $W + jets$	±10 – 30%	±10%	±10 – 30%	±30%
Norm. Z: l fakes τ_h	±20 – 100%	±20 – 30%	±20 – 100%	±30%
Norm. Z: jet fakes τ_h	±20%	±20%	±20%	±30%

Theory Uncertainties (SM)		Propagation into Limit Calculation		
Uncertainty	Uncert.	*0-Jet*	*Boost*	*VBF*
PDF (*)	-	±2 – 8%	±2 – 8%	±2 – 8%
$\mu_r/\mu_f(gg \to H)$ (*)	-	±8%	±10%	±30%
$\mu_r/\mu_f(qq \to H)$ (*)	-	±3.5%	±4%	±10%
$\mu_r/\mu_f(qq \to VH)$ (*)	-	±4%	±4%	±4%
UE & PS (*)	-	∓4%	±4%	±4%

Conclusions

- Presented most recent results on search for SM H → $b\bar{b}$ at CMS
  → Improved VH analysis on 2011+2012 data
  → First ttH analysis on 2011 data

- Mild excess in VH analysis, exp(obs) limit at m_H(125)=1.6(2.)
  will likely reach Standard Model sensitivity by end of 2012 !

- No excess in ttH, exp(obs) limit at m_H(125) = 4.6(3.8)
  additional information from this channel on Higgs properties